Phase Equilibria and Critical Properties of Carbon Dioxide Obtained by Molecular Dynamics Simulations

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Christoph Bratschi
aus
Safnern (BE)

Basel, den 21. März 2005

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. H. Huber und Prof. Dr. M. Meuwly

Basel, den 10. April 2005

Prof. Dr. Hans-Jakob Wirz
Dekan

Contents

List of Figures

List of Tables

List of Symbols

Symbols and Constants

A	Helmholtz free energy
$\boldsymbol{a}$	Acceleration
BL	Box length
D	Spatial dimensions
$D(V)$	Spatial domain
E	Energy
$\boldsymbol{F}$	Molecular force
f	Probability density function
$\boldsymbol{f}$	Force
G	Gibbs free energy
g	Degrees of freedom
$\sqrt{g}$	Metric determinant factor
$\mathcal{H}$	Hamiltonian
h	Planck's constant
$\mathcal{K}$	Kinetic energy
k	Spring constant
k_B	Boltzmann's constant
L	Gibbs ensemble subsystems
$\mathcal{L}$	Lagrangian
M	Molecular mass
m	Mass
N	Number of molecules
P	Pressure
P_E	External pressure
$\boldsymbol{p}$	Generalized momenta
p_s	Conjugated momenta of s
p_ν	Conjugated momenta of ν
Q	Thermostat mass
Q_P	Barostat mass
Q_E	Exchange mass
$\boldsymbol{q}$	Generalized coordinate
R	Universal gas constant
$\boldsymbol{R}$	Center of mass coordinate
$\boldsymbol{r}$	Scaled coordinate (range [0,1])
s	Nosé's additional degree of freedom
S	Entropy/Action

T	Temperature
T_E	External temperature
t	Time
U	Internal energy
V	Volume
V_m	Molar volume
V_0	Arbitrary volume
$\boldsymbol{v}$	Velocity
v_0	Initial exchange velocity
$\mathcal{W}$	Virial
w	Primitive function of κ
$\wp$	Probability
X	Degrees of freedom

Greek Symbols

β	Inverse temperature: $\frac{1}{k_B T}$
δ	Dirac delta function
δt	Time step
ϵ	Barostat position
η	Thermostat position
	Maximal distance increase in the particle exchange
$\boldsymbol{\eta}$	Phase space vector
κ	Compressibility factor
Λ	Thermal de Broglie wavelength: $\sqrt{\frac{h^2}{2\,\pi\,m\,k_B\,T}}$
μ	Chemical potential
ν	Exchange position
ν_0	Initial exchange position
ξ	Thermodynamic friction coefficient
ξ_P	Strain rate
ρ	Density
τ	Oscillation period
Φ	Potential function
Ω	Partition function

Abbreviations

ACF	Time auto-correlation function
BBV	Bock-Bich-Vogel potential
GEMC	Gibbs ensemble Monte Carlo
GEMD	Gibbs ensemble molecular dynamics
MC	Monte Carlo
MD	Molecular dynamics
NH	Nosé-Hoover
NHC	Nosé-Hoover chain
NPT	Isothermal-isobaric ensemble
NVE	Microcanonical ensemble
NVT	Canonical ensemble
SAPT	Symmetry-adapted perturbation theory
SimMol	Name of the simulation program[11]

Mathematical Conventions

Scalar:	v_x
Vector:	$\boldsymbol{v}$
Tensor:	M
Cross product:	$\boldsymbol{v_1} \times \boldsymbol{v_2}$
Scalar product:	$\boldsymbol{v_1} \cdot \boldsymbol{v_2}$
Tensor product:	$\boldsymbol{v_1} \otimes \boldsymbol{v_2}$
Determinant:	$\mathrm{Det}(\mathsf{M})$
Trace:	$\mathrm{Tr}(\mathsf{M})$
Phase space gradient:	$\boldsymbol{\nabla_\eta}$
Single time derivative:	$\dot{\boldsymbol{q}}$
Double time derivative:	$\ddot{\boldsymbol{q}}$
Average value:	$<>$

1. Introduction

The initial goal of this work was to study the quality of two recently published CO_2 *ab initio* pair potentials. Carbon dioxide is an important solvent in the pharmaceutical and chemical process industry and is needed in large quantities for the production of methanol and urea. If CO_2 is dissolved in oil the oil becomes less viscous and can be more easily extracted from the well. Supercritical CO_2 has a similar solvent power as petrol and has the ability to dissolve fluorinated organic compounds. At ambient conditions CO_2 is gaseous and can be easily separated and recycled from liquid and solid substances. CO_2 is a replacement for more toxic chemicals. In the atmosphere it is a greenhouse gas responsible for global warming. Therefore it is of interest to predict properties by a pure theoretical computational model.

First checks using molecular dynamics simulations in the liquid phase showed that the best CO_2 potential delivers accurate results close to experimental values. An open question of all CO_2 *ab initio* potentials was where the critical point is and how well the potentials reproduce liquid-vapor equilibria and supercritical states. To obtain the liquid-vapor phase equilibria, the Gibbs ensemble Monte Carlo method had to be extended to molecular dynamics. A new algorithm was developed using state of the art non-Hamiltonian molecular dynamics techniques. It is the first time that the Gibbs ensemble molecular dynamics method with this kind of algorithm is used for a molecular system as CO_2. The results show that the new algorithm works and the best CO_2 potential delivers values close to the experimental ones.

Supercritical fluids exhibit interesting phenomena in the vicinity of the critical point. These phenomena are independent of the substance. Supercritical fluids have the ability to fill any volume like a gas and dissolve materials like a liquid. By small changes in temperature or pressure the fluid properties change dramatically and the desired properties can be tuned. The diffusion rates are high and the viscosity is low, both attributes accelerate diffusion limited reactions. Supercritical CO_2 is used for the decaffination of coffee and tea. A high industrial potential lies in polymerization processes. The supercritical properties were studied with methods taken from the literature.

General Introduction

For roughly sixty years, computer simulations have been used to solve microscopic, macroscopic and general problems. The main progress was the ability to define a mathematical problem, solve it at an incredible speed compared to a human and to store and reuse existing results. Due to the rapid increase in computational power, the models were enhanced and knowledge grew. Since one always wants to get the best accuracy, the models were extended to the limits of the computers.

In molecular dynamics, Lennard-Jones potentials, together with Coulomb interactions for charged species, have a long tradition. It is well known that a 10-6 or exp-6 potential would improve the repulsive part, compared to the standard 12-6 potential. But this would increase the computational cost and the potential is used as it is. To obtain more accurate values, the attention is paid to the parameters. The parameters are optimized to reproduce certain empirical properties in a limited phase space region. Therefore different sets of parameters for different conditions exist.

A different strategy is to use *ab initio* methods. The fundamental interactions are calculated by mathematically well defined methods based on the axioms of quantum mechanics. The results are always approximations of the real values (with a few exceptions) but improvements in the model and additional computer power refine the results. If a converging method was used, the limits of the parameters can be estimated by extrapolations. The accuracy of *ab initio* molecular dynamics is given by the classical equations of motion and the quality of the potential. Instead of using a precalculated potential, the interactions can be calculated on the fly, by quantum calculations. At each time step the energies and forces of all atoms have to be calculated. This limits the system size and the calculation depth. Today, this means Car-Parrinello calculations using the Density Functional Theory which is often used as a non-*ab initio* method by introducing empirical parameters. And there exists the so far unsolved problem of including the dispersion in the functional.

Chapters

The first chapters are introductions to mechanics, molecular dynamics and Monte Carlo methods. This is followed by an introduction to extended molecular dynamics ensembles. The next chapter introduces the Gibbs ensemble method and its molecular dynamics implementation. Results for CO_2 are then presented. The last chapter concerns the supercritical fluid behavior in the vicinity of the critical point. Additional details of the simulation program are given in the appendices.

2. Classical and Statistical Mechanics

In this chapter, a short introduction into the concepts of classical and statistical mechanics is given. The equations shown here will be used in the following chapters. The overview starts with Newton's empirical laws, continues with the Lagrangian and Hamiltonian formulations, and finally the non-Hamiltonian theory completes the classical mechanics chapter.

The statistical mechanics part is a brief overview of common concepts which intersect with this work. More details are embedded into subsequent chapters.

2.1. Newtonian, Lagrangian and Hamiltonian Mechanics

Nowadays, several different formulations of classical mechanics exist. The first formulation was published by Isaac Newton[96] in the year 1687. The empirical laws of universal motion are:

- Law of inertia/Galileo's principle: *Every body continues in its state of rest, or of uniform motion in a right (straight) line, unless it is compelled to change that state by forces impressed upon it.*

$$\dot{\boldsymbol{v}} = 0 \tag{2.1}$$

- Fundamental law of dynamics: *The time rate of change in momentum is proportional to the net force acting on the object and takes place in the direction of the force.*

$$\boldsymbol{f} = m\,\boldsymbol{a} \tag{2.2}$$

  In differential form with generalized momentum and coordinate:

$$\dot{\boldsymbol{p}} = m\,\ddot{\boldsymbol{q}} \tag{2.3}$$

- Law of reciprocal actions: *Whenever one body exerts force upon a second body, the second body exerts an equal and opposite force upon the first body.*

$$\boldsymbol{f}_{ab} = -\boldsymbol{f}_{ba} \tag{2.4}$$

These laws are sufficient for standard molecular dynamics simulations. More advanced formulations are used to describe more complex systems. Two equivalent theories are the Hamiltonian and Lagrangian formulation of mechanics.

Example system

In the following sections an example system with the following properties is used. The three dimensional system has N particles. No external fields or forces are present. The interactions between the particles are given by the potential $\Phi(\boldsymbol{q})$. A classical behavior is assumed, quantum and relativistic effects are neglected.

2.1.1. Lagrangian Mechanics

The Lagrangian formulation of classical mechanics[24, Appendix A] is based on the variational principle

$$S = \int_{t_b}^{t_e} \mathrm{d}t \, \mathcal{L}\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) \tag{2.5}$$

where the action S is an extremum, usually a minimum. $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ is called the Lagrangian. The Lagrangian is a function of generalized coordinates and generalized velocities. The Lagrangian for the example system is the kinetic energy $\mathcal{K}$ minus the potential energy Φ of all particles.

$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{K}(\dot{\boldsymbol{q}}) - \Phi(\boldsymbol{q}) \tag{2.6}$$

Free particle

To prove that the Lagrangian formulation is equivalent to Newton's, the simplest case is considered: one particle in a one dimensional system without an external force. The particle moves from x_b to x_e in the time t_e - t_b.

The Lagrangian of the system is:

$$\mathcal{L}(\dot{x}) = \frac{1}{2}\, m\, \dot{x}^2(t) \tag{2.7}$$

The average velocity of the particle is:

$$v_{av} = \frac{x_e - x_b}{t_e - t_b} \tag{2.8}$$

The path with constant average velocity is called $\bar{x}(t)$. The true trajectory and velocity are defined as:

$$\begin{aligned} x(t) &= \bar{x}(t) + \eta(t) && (2.9)\\ v(t) &= \dot{x}(t) && (2.10)\\ &= v_{av} + \dot{\eta}(t) \end{aligned}$$

$\eta(t)$ is the non-linear position function, its value is so far unknown. From the boundary conditions $\eta(t_b) = 0$ and $\eta(t_e) = 0$ it follows that the integral along the path of the time derivative of $\eta(t)$ has to be zero.

$$\int_{t_b}^{t_e} \mathrm{d}t\, \dot{\eta}(t) = 0 \tag{2.11}$$

It follows that the action of the true path is

$$\begin{aligned} \mathcal{L}(t) &= \frac{1}{2}\, m\, (v_{av} + \dot{\eta}(t))^2 && (2.12)\\ S &= \frac{1}{2}\, m \int_{t_b}^{t_e} \mathrm{d}t\, [v_{av} + \dot{\eta}(t)]^2 && (2.13)\\ &= \bar{S} + \frac{1}{2}\, m \int_{t_b}^{t_e} \mathrm{d}t\, \dot{\eta}^2(t) && (2.14) \end{aligned}$$

where the first term $\bar{S}$ is the average action along the shortest path. The second term is always greater than or equal to zero. Therefore setting $\dot{\eta}$ to zero gives the minimum action. The particle always has the average velocity, this is equal to Newton's first principle.

Particle in a potential

The second case uses an external potential. The Lagrangian and action of this system are:

$$\mathcal{L} = \frac{1}{2} m \dot{x}^2(t) - \Phi\left(x(t)\right) \tag{2.15}$$

$$S = \int_{t_b}^{t_e} \mathrm{d}t \left[\frac{1}{2} m \dot{x}^2(t) - \Phi\left(x(t)\right)\right] \tag{2.16}$$

The true path formalism, used for the free particle, is introduced again.

$$S = \int_{t_b}^{t_e} \mathrm{d}t \left[\frac{1}{2} m \left(\dot{\bar{x}}(t) + \dot{\eta}(t)\right)^2 - \Phi\left[\bar{x}(t) + \eta(t)\right]\right] \tag{2.17}$$

It is assumed that non-linear values, obtained by the functions $\eta(t)$ and $\dot{\eta}(t)$, are small. Quadratic terms are neglected and a functional expansion can be obtained for the potential

$$S \approx \int_{t_b}^{t_e} \mathrm{d}t \frac{1}{2} m \left[\dot{\bar{x}}^2(t) + 2\,\dot{\bar{x}}(t)\,\dot{\eta}(t)\right] - \left[\Phi\left(\bar{x}(t)\right) + \frac{\partial\,\Phi\left(\bar{x}(t)\right)}{\partial\,x}\,\eta(t)\right] \tag{2.18}$$

$$= \bar{S} + \int_{t_b}^{t_e} \mathrm{d}t \left[m\,\dot{\bar{x}}(t)\,\dot{\eta}(t) - \frac{\partial\,\Phi\left(\bar{x}(t)\right)}{\partial\,x}\,\eta(t)\right] \tag{2.19}$$

where $\bar{S}$ is the average action. Partial integration gives the final result:

$$S = \bar{S} + m\,\dot{\bar{x}}(t)\,\eta(t)\Big|_{t_b}^{t_e} - \int_{t_b}^{t_e} \mathrm{d}t \left[m\,\ddot{\bar{x}}(t) + \frac{\partial\,\Phi\left(\bar{x}(t)\right)}{\partial\,x}\right]\eta(t) \tag{2.20}$$

$$= \bar{S} - \int_{t_b}^{t_e} \mathrm{d}t \left[m\,\ddot{\bar{x}}(t) + \frac{\partial\,\Phi\left(\bar{x}(t)\right)}{\partial\,x}\right]\eta(t) \tag{2.21}$$

One obtains the extremal action if the integrand vanishes for arbitrary $\eta(t)$. This condition is satisfied with the relation

$$m\,\ddot{\bar{x}}(t) = -\frac{\partial\,\Phi\,(\bar{x}(t))}{\partial\,x} \tag{2.22}$$

which is nothing other than Newton's equation of motion.

Properties

Generalized momenta $\boldsymbol{p}_i$ and forces $\dot{\boldsymbol{p}}_i$ are obtained from the relations:

$$\boldsymbol{p}_i \equiv \frac{\partial\,\mathcal{L}\,(\boldsymbol{q},\dot{\boldsymbol{q}})}{\partial\,\dot{\boldsymbol{q}}_i} \tag{2.23}$$

$$\dot{\boldsymbol{p}}_i \equiv \frac{\partial\,\mathcal{L}\,(\boldsymbol{q},\dot{\boldsymbol{q}})}{\partial\,\boldsymbol{q}_i} \tag{2.24}$$

Application to the example system gives:

$$\boldsymbol{q} = (\boldsymbol{q}_1,\boldsymbol{q}_2,..,\boldsymbol{q}_N) \tag{2.25}$$

$$\dot{\boldsymbol{q}} = (\dot{\boldsymbol{q}}_1,\dot{\boldsymbol{q}}_2,..,\dot{\boldsymbol{q}}_N) \tag{2.26}$$

$$\mathcal{K}(\dot{\boldsymbol{q}}) = \sum_{i=1}^{N}\frac{1}{2}\,m_i\,\dot{\boldsymbol{q}}_i^2 \tag{2.27}$$

$$\Phi(\boldsymbol{q}) = \sum_i \Phi_1\,(\boldsymbol{q_i}) + \sum_i\sum_{j>1}\Phi_2\,(\boldsymbol{q_i},\boldsymbol{q_j}) + \cdots \tag{2.28}$$

$$\mathcal{L}(\boldsymbol{q},\dot{\boldsymbol{q}}) = \mathcal{K}(\dot{\boldsymbol{q}}) - \Phi(\boldsymbol{q}) \tag{2.29}$$

$$\boldsymbol{p}_i = m_i\,\dot{\boldsymbol{q}}_i \tag{2.30}$$

$$\dot{\boldsymbol{p}}_i = -\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}_i} \tag{2.31}$$

$$= \boldsymbol{f}_i \tag{2.32}$$

$$\ddot{\boldsymbol{q}}_i = -\frac{1}{m_i}\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}_i} \tag{2.33}$$

Where the last equation, Newton's equation of motion, is a second order differential equation.

The general potential defined in equation (2.28) includes the intramolecular potential Φ_1, the pair potential Φ_2 (see section 7.1 on page 72) and possible higher terms like three-body interactions.

2.1.2. Hamiltonian Mechanics

The Hamiltonian[3, 24] $\mathcal{H}(\boldsymbol{q}, \boldsymbol{p}, t)$ is a function of generalized coordinates, generalized momenta and the time. The Hamiltonian can be obtained by Legendre's transformation from the Lagrangian:

$$\mathcal{H}(\boldsymbol{q}, \boldsymbol{p}, t) \equiv \boldsymbol{p}\dot{\boldsymbol{q}} - \mathcal{L}\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) \tag{2.34}$$

The total differential of the Hamiltonian, the left side of Legendre's equation, is:

$$\mathrm{d}\mathcal{H}(\boldsymbol{q}, \boldsymbol{p}, t) = \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}}\,\mathrm{d}\boldsymbol{p} + \frac{\partial \mathcal{H}}{\partial \boldsymbol{q}}\,\mathrm{d}\boldsymbol{q} + \frac{\partial \mathcal{H}}{\partial t}\,\mathrm{d}t \tag{2.35}$$

The total differential of the right side of Legendre's transformation is:

$$\begin{aligned}
\mathrm{d}\mathcal{H}(\boldsymbol{q}, \boldsymbol{p}, t) &= \mathrm{d}\left(\boldsymbol{p}\dot{\boldsymbol{q}}\right) - \mathrm{d}\mathcal{L}\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) && (2.36)\\
&= \boldsymbol{p}\,\mathrm{d}\dot{\boldsymbol{q}} + \dot{\boldsymbol{q}}\,\mathrm{d}\boldsymbol{p} - \left[\frac{\partial \mathcal{L}}{\partial \boldsymbol{q}}\,\mathrm{d}\boldsymbol{q} + \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}\,\mathrm{d}\dot{\boldsymbol{q}} + \frac{\partial \mathcal{L}}{\partial t}\,\mathrm{d}t\right] && (2.37)\\
&= \boldsymbol{p}\,\mathrm{d}\dot{\boldsymbol{q}} + \dot{\boldsymbol{q}}\,\mathrm{d}\boldsymbol{p} - \dot{\boldsymbol{p}}\,\mathrm{d}\boldsymbol{q} - \boldsymbol{p}\,\mathrm{d}\dot{\boldsymbol{q}} - \frac{\partial \mathcal{L}}{\partial t}\,\mathrm{d}t && (2.38)\\
&= \dot{\boldsymbol{q}}\,\mathrm{d}\boldsymbol{p} - \dot{\boldsymbol{p}}\,\mathrm{d}\boldsymbol{q} - \frac{\partial \mathcal{L}}{\partial t}\,\mathrm{d}t && (2.39)
\end{aligned}$$

The equations of motion, Hamilton's equations, are obtained by combining the equations (2.35) and (2.39):

$$\begin{aligned}
\dot{\boldsymbol{q}}_i &= \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_i} && (2.40)\\
\dot{\boldsymbol{p}}_i &= -\frac{\partial \mathcal{H}}{\partial \boldsymbol{q}_i} && (2.41)\\
\frac{\partial \mathcal{H}}{\partial t} &= -\frac{\partial \mathcal{L}}{\partial t} && (2.42)
\end{aligned}$$

If the Lagrangian is not a function of time, then the Hamiltonian itself is a conserved quantity, namely the total energy:

$$\dot{\mathcal{H}}(\boldsymbol{q}, \boldsymbol{p}) = \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}}\,\mathrm{d}\dot{\boldsymbol{p}} + \frac{\partial \mathcal{H}}{\partial \boldsymbol{q}}\,\mathrm{d}\dot{\boldsymbol{q}} \tag{2.43}$$

$$= -\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}}\frac{\partial \mathcal{H}}{\partial \boldsymbol{q}} + \frac{\partial \mathcal{H}}{\partial \boldsymbol{q}}\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}} \tag{2.44}$$

$$\equiv 0 \tag{2.45}$$

Properties

The value of the Hamiltonian of a closed system is the total energy. Other conserved quantities, if no external field is applied, are the total momentum and the angular momentum. The potential is independent of the time and momenta and a function of the coordinates. In molecular dynamics simulations periodic boundary conditions are usually used[3, p. 73]. In this case the angular momentum is no longer a conserved quantity. Following Newton's third law, the total force of the system is by definition zero and therefore the total linear momentum is a conserved quantity (equations (2.46) and (2.47)). The equations of motion obtained are time reversible, *i.e.* the same trajectory is obtained again if the direction of the time (sign) is changed. In a macroscopic system this would be a violation of the second law of thermodynamics[1] but this is an important property of a microscopic system[39]. Hoover analyzed the macroscopic irreversibility and wrote an interesting formulation of the second law: "the future is more nearly predictable than is the past". Recently Wang *et al.*[94] were able to measure these violations, the reversibility, in small systems.

$$\sum_i^N \dot{\boldsymbol{p}}_i = \mathbf{0} \tag{2.46}$$

$$\sum_i^N \boldsymbol{p}_i = \mathbf{const} \tag{2.47}$$

The example system in Hamiltonian notation is shown below. Each particle has a generalized momentum $\boldsymbol{p}_i$ and a generalized coordinate $\boldsymbol{q}_i$:

$$\boldsymbol{q} = (\boldsymbol{q}_1, \boldsymbol{q}_2, .., \boldsymbol{q}_N) \tag{2.48}$$

$$\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2, .., \boldsymbol{p}_N) \tag{2.49}$$

The Hamiltonian $\mathcal{H}$ is defined as the sum of the kinetic and potential energy.

$$\mathcal{H}(\boldsymbol{q}, \boldsymbol{p}) = \mathcal{K}(\boldsymbol{p}) + \Phi(\boldsymbol{q}) \tag{2.50}$$

$$\mathcal{K} = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2\, m_i} \tag{2.51}$$

[1] The law of entropy.

$$\Phi = \sum_i \Phi_1(\boldsymbol{q}_i) + \sum_i \sum_{j>1} \Phi_2\left(\boldsymbol{q}_i, \boldsymbol{q}_j\right) + \cdots \tag{2.52}$$

The velocities and forces, the equations of motion, are calculated with equations (2.40) and (2.41). Two first order differential equations are the result, whereas in the Lagrangian case, one first order and a second order equation were obtained.

Symplectic Condition

The phase space area is a preserved quantity of any canonical transformation. This is called the symplectic condition[24, p. 491].

For a given phase space vector of a one dimensional system, containing all coordinates and momenta, the size of the vector is $2\,N$.

$$\boldsymbol{\eta} = (q_1 \cdots q_N, p_1 \cdots p_N) \tag{2.53}$$
$$= (\eta_1 \cdots \eta_{2N}) \tag{2.54}$$

The time derivative of the phase space vector is

$$\dot{\boldsymbol{\eta}} = \omega \frac{\partial \mathcal{H}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \tag{2.55}$$

where ω is an antisymmetric matrix. With the time derivatives given by the equations (2.40) and (2.41), the matrix ω of the example system has the following structure:

$$\omega = \left(\begin{array}{ccc|ccc} & & & 1 & & 0 \\ & 0 & & & \ddots & \\ & & & 0 & & 1 \\ \hline -1 & & 0 & & & \\ & \ddots & & & 0 & \\ 0 & & -1 & & & \end{array}\right) \tag{2.56}$$

The initial phase space vector $\boldsymbol{\eta}$ is now transformed to the new vector $\boldsymbol{\xi}$. The transformation of the time derivatives is

$$\dot{\boldsymbol{\xi}} = \mathsf{M}\,\dot{\boldsymbol{\eta}} \tag{2.57}$$

$$= \mathsf{M}\,\omega\,\frac{\partial\,\mathcal{H}(\boldsymbol{\eta})}{\partial\,\boldsymbol{\eta}} \tag{2.58}$$

where M is the Jacobian matrix with elements $\mathsf{M}_{ij} = \frac{\partial\,\xi_i}{\partial\,\eta_j}$.

Considering the inverse transformation, the following relation can be used

$$\frac{\partial\,\mathcal{H}(\boldsymbol{\eta})}{\partial\,\eta_i} = \sum_j \frac{\partial\,\mathcal{H}(\boldsymbol{\xi})}{\partial\,\xi_j}\,\frac{\partial\,\xi_j}{\partial\,\eta_i} \tag{2.59}$$

$$= \sum_j \frac{\partial\,\mathcal{H}(\boldsymbol{\xi})}{\partial\,\xi_j}\,\mathsf{M}^{\mathsf{T}}_{ij} \tag{2.60}$$

where M^{T} is the transposed Jacobian matrix.

Substitution of equation (2.60) into (2.58) leads to:

$$\dot{\boldsymbol{\xi}} = \mathsf{M}\,\omega\,\mathsf{M}^{\mathsf{T}}\,\frac{\partial\,\mathcal{H}(\boldsymbol{\xi})}{\partial\,\boldsymbol{\xi}} \tag{2.61}$$

The transformation is canonical if the symplectic condition is satisfied:

$$\mathsf{M}\,\omega\,\mathsf{M}^{\mathsf{T}} = \omega \tag{2.62}$$

In this case the matrix M is called a symplectic matrix.

Liouville's Theorem

Different states of the system, given by a Hamiltonian, are defined by a phase space vector $\boldsymbol{\eta}$. The phase space distribution function $f(\boldsymbol{\eta}, t)$ gives the probability of a phase space vector. An expectation value $A(\boldsymbol{\eta})$ is obtained by integration over the whole phase space and by normalization:

$$< A >= \frac{\int \mathrm{d}\boldsymbol{\eta}\,A(\boldsymbol{\eta})\,f(\boldsymbol{\eta})}{\int \mathrm{d}\boldsymbol{\eta}\,f(\boldsymbol{\eta})} \tag{2.63}$$

An important relation is that the phase space probability density $f(\boldsymbol{\eta}, t)$ and the phase space volume $\mathrm{d}\boldsymbol{\eta}$ are time independent quantities.

$$\frac{\partial\,[f(\boldsymbol{\eta}, t)\,\mathrm{d}\boldsymbol{\eta}]}{\partial\, t} \;=\; \mathrm{d}\boldsymbol{\eta}\,\frac{\partial\, f}{\partial\, t} \tag{2.64}$$

That means that the total probability of a phase space vector moving through the phase space is conserved. The overall probability flux is zero (see continuity equation (2.65)).

$$\frac{\partial\, f}{\partial\, t} + \boldsymbol{\nabla}_{\boldsymbol{\eta}} \cdot (f\,\dot{\boldsymbol{\eta}}) \;=\; 0 \tag{2.65}$$

If the Hamiltonian phase space is inserted in the continuity equation (2.65) then one obtains equation (2.66) which is called Liouville's theorem[38]. Liouville's equation is equal to the total differential of the phase space probability density divided by $\mathrm{d}t$.

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\partial\, f}{\partial\, t} + \dot{\boldsymbol{\eta}} \cdot \boldsymbol{\nabla}_{\boldsymbol{\eta}} f \\ &= 0 \end{aligned} \tag{2.66}$$

A consequence of equation (2.64) is that an invariant phase space measure $\mathrm{d}\boldsymbol{\eta}$, called Liouville measure, exists (equation (2.86)). This quantity is the multidimensional "volume" of the phase space. As shown in the symplectic condition chapter, Hamiltonian dynamics is area preserving.

Partition Function

Each dynamic system has conservation laws. The Hamiltonian, which is equal to the total energy, is for example a conserved quantity. In general, each conserved quantity λ_i is defined by its value and a function Λ_i which determines the quantity:

$$\Lambda_i(\boldsymbol{\eta}) \;=\; \lambda_i \tag{2.67}$$

$$\dot{\Lambda}_i(\boldsymbol{\eta}) \;=\; 0 \tag{2.68}$$

In the Hamiltonian case is $\Lambda = \mathcal{H}(\boldsymbol{\eta})$ and $\lambda = E$.

The phase space distribution function is given by products of Dirac delta functions

$$f(\boldsymbol{\eta}) = \prod_{i=1}^{n_c} \delta\left(\Lambda_i(\boldsymbol{\eta}) - \lambda_i\right) \tag{2.69}$$

where n_c is the number of conserved values.

The calculation of average values (equation (2.63)) scans the complete phase space. Only values obeying the conservation laws contribute to the left hand side of the equation. Geometrically this corresponds to the intersection of the hypersurface given by the delta functions with the infinite phase space.

The number of microscopic states Ω, called the partition function, is the integral of the phase space distribution function.

$$\Omega = \int \mathrm{d}\boldsymbol{\eta}\, f(\boldsymbol{\eta}) \tag{2.70}$$

The Gibbs postulate says that all microscopic states have the same probability. The delta functions satisfy this condition. The contribution is either one or zero (or a constant prefactor for each state).

Insertion into equation (2.63) leads to the following expression for phase space averages

$$< A >= \frac{1}{\Omega} \int \mathrm{d}\boldsymbol{\eta}\, A(\boldsymbol{\eta})\, f(\boldsymbol{\eta}). \tag{2.71}$$

Ergodic Hypothesis

The ergodic hypothesis states that ensemble averages are equal to time averages. The ensemble average given by equation (2.71) is the exact value. The phase space vectors obtained from a molecular dynamics trajectory can only sample a subset of all possible states. That means ergodicity is important in order to be able to calculate any properties at all. A consequence of working in an ergodic system is that all averages are independent of the initial state.

$$< A >_{\text{ensemble}} = \bar{A}_{\text{time}} \tag{2.72}$$

$$\bar{A} = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}t\, A\left(\boldsymbol{\eta}(t)\right) \tag{2.73}$$

If the time integral is infinite, all microscopic states will be counted at least once, if the motion is chaotic enough[39]. In practice an ideal time length does not exist and many factors play a role. In fact, ergodicity is assumed for most systems, but can only be proved for a small number of systems, mostly simple single particle systems.

2.2. Statistical Mechanics

Statistical thermodynamics and statistical mechanics are both theories to describe macroscopic systems. Macroscopic means that the number of particles is of the order of molar quantities, about 10^{23} particles. The fundamental physical theories to describe general systems are classical mechanics for the dynamics and electrostatic interactions between particles. The macroscopic system possesses several invariant properties, for example the number of particles, the temperature or pressure. Additional properties or expectation values are obtained by the relations known from statistical thermodynamics. Most properties can be measured in experiments.

To be able to describe macroscopic systems, probability distribution functions and statistical methods are used. Each macroscopic ensemble consists of a collection of all microscopic states. The partition function is the integration of the probability distribution function, giving the total number of microscopic states[2].

2.2.1. Microcanonical Ensemble

The number of particles N, the volume V and the total energy E are constant in the microcanonical ensemble. This is the default ensemble used in molecular dynamics simulations.

The probability distribution function is proportional to[3]:

$$f_{\mathrm{NVE}} \propto \delta\left(\mathcal{H}(\boldsymbol{p},\boldsymbol{q}) - E\right) \tag{2.74}$$

The partition function of a three dimensional microcanonical system is defined as:

$$\Omega_{\mathrm{NVE}} = \frac{1}{N!\,h^{3N}} \int \mathrm{d}^N\boldsymbol{p} \int \mathrm{d}^N\boldsymbol{q}\, \delta\left(\mathcal{H}(\boldsymbol{p},\boldsymbol{q}) - E\right) \tag{2.75}$$

The prefactor $1/N!\,h^{3N}$ is given by the requirement that the entropy of an ideal gas is zero and the indistinguishability factor $1/N!$

[2] See Hamiltonian chapter on the previous pages.

$$S = k_B \ln\left(\Omega_{\mathrm{NVE}}\right). \tag{2.76}$$

2.2.2. Canonical Ensemble

The number of particles N, the volume V and the temperature T are constant in the canonical ensemble. The energy is not constant. This ensemble is usually used in Monte Carlo calculations.

The probability distribution function is proportional to:

$$f_{\mathrm{NVT}} \propto \mathrm{e}^{-\beta\,\mathcal{H}(\boldsymbol{p},\boldsymbol{q})} \tag{2.77}$$

The partition function is defined as:

$$\Omega_{\mathrm{NVT}} = \frac{1}{N!\,h^{3N}} \int \mathrm{d}^N\boldsymbol{p} \int \mathrm{d}^N\boldsymbol{q}\, \mathrm{e}^{-\beta\,\mathcal{H}(\boldsymbol{p},\boldsymbol{q})} \tag{2.78}$$

The appropriate thermodynamic function is the Helmholtz free energy

$$A = -k_B\,T \ln\left(\Omega_{\mathrm{NVT}}\right). \tag{2.79}$$

2.2.3. Isothermal-Isobaric Ensemble

The number of particles N, the pressure P and the temperature T are constant in the isothermal-isobaric ensemble. This ensemble is useful to draw comparisons with experimental values.

The probability distribution function is proportional to:

$$f_{\mathrm{NPT}} \propto \mathrm{e}^{-\beta\,(\mathcal{H}(\boldsymbol{p},\boldsymbol{q})+P\,V)} \tag{2.80}$$

The partition function is defined as:

$$\Omega_{\mathrm{NPT}} = \frac{1}{N!\, h^{3N}\, V_0} \int_0^\infty \mathrm{d}V\, \mathrm{e}^{-\beta P V} \int \mathrm{d}^N \boldsymbol{p} \int_{D(V)} \mathrm{d}^N \boldsymbol{q}\, \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{p},\boldsymbol{q})} \tag{2.81}$$

The appropriate thermodynamic function is the Gibbs free energy

$$G = -k_B T\, \ln\left(\Omega_{\mathrm{NPT}}\right). \tag{2.82}$$

2.3. Non-Hamiltonian Mechanics

To construct ensembles other than the microcanonical one, some properties have to be fixed, for example the temperature in the NVT ensemble. The easiest way to do this is to extend a Lagrangian. In general, extended Lagrangians cannot be transformed into a Hamiltonian form and the revealing equations of motion are called non-Hamiltonian. Therefore a new theory is needed which connects the non-Hamiltonian dynamics to statistical mechanics.

In 1999 Tuckerman *et al.* presented a consistent classical statistical mechanical theory for non-Hamiltonian dynamical systems[88]. In this chapter this theory will be summarized using the original paper and additional literature[86, 87, 41, 24, 78]. The non-Hamiltonian theory unites several incomplete theories published earlier for special cases. The theory is accepted and very useful for extended systems which could not be completely described by earlier theories.

It was shown before (see section 2.1.2) that Hamiltonian dynamics satisfies the symplectic condition, a volume element of the phase space has always the same volume. Solving the equations of motion is nothing other than a coordinate transformation

$$\mathrm{d}\boldsymbol{\eta}_t = J\left(\boldsymbol{\eta}_t; \boldsymbol{\eta}_0\right)\, \mathrm{d}\boldsymbol{\eta}_0 \tag{2.83}$$

where J is the determinant of the Jacobian matrix M of the transformation.

$$J\left(\boldsymbol{\eta}_t; \boldsymbol{\eta}_0\right) = \mathrm{Det}(\mathsf{M}) \tag{2.84}$$

$$M_{ij} = \frac{\partial\, \eta_{i,t}}{\partial\, \eta_{j,0}} \tag{2.85}$$

In the case of a canonical transformation, J is ± 1[3]. The Hamiltonian phase space is incompressible, in other words: the phase space is flat. The preserved area (equation (2.86)), called the Liouville measure, and its integral (equation (2.87)) are a constant of motion.

[3] A simple Ansatz for a proof is given in [53].

$$\prod_{i=1}^{N} \mathrm{d}\boldsymbol{p}_i \mathrm{d}\boldsymbol{q}_i \tag{2.86}$$

$$\prod_{i=1}^{N} \int \mathrm{d}\boldsymbol{p}_i \mathrm{d}\boldsymbol{q}_i \tag{2.87}$$

On the other hand, if the time dependence of the phase space vector of a non-Hamiltonian system is analyzed, a Jacobian which in general is different to one is observed. This means that the phase space is compressible and no longer flat. This fact has to be included into the Liouville theorem and other relations to generate an incompressible partition function.

If the Jacobian matrix M is triangular[4], the determinant is the product of all diagonal elements. Equation (2.84) can be written as:

$$\begin{aligned} J\left(\boldsymbol{\eta}_t; \boldsymbol{\eta}_0\right) &= \mathrm{e}^{\mathrm{Tr}(\ln \mathsf{M})} && (2.88) \\ &= \prod_{i=1}^{N} M_{ii} && (2.89) \end{aligned}$$

The time derivative of the Jacobian is

$$\begin{aligned} \dot{J}\left(\boldsymbol{\eta}_t; \boldsymbol{\eta}_0\right) &= J\left(\boldsymbol{\eta}_t; \boldsymbol{\eta}_0\right) \mathrm{Tr}\left(\mathsf{M}^{-1}\dot{\mathsf{M}}\right) && (2.90) \\ &= J\left(\boldsymbol{\eta}_t; \boldsymbol{\eta}_0\right) \sum_{i=1}^{N} M_{ii}^{-1} \dot{M}_{ii} && (2.91) \end{aligned}$$

where the matrix elements are:

$$M_{ij}^{-1} = \frac{\partial\, \eta_{j,0}}{\partial\, \eta_{i,t}} \tag{2.92}$$

$$\dot{M}_{ij} = \frac{\partial\, \dot{\eta}_{i,t}}{\partial\, \eta_{j,0}} \tag{2.93}$$

Including these relations into equation (2.91) leads to:

[4] To decompose any matrix into a triangular matrix the matrix has to be invertible. Therefore no further restrictions are introduced.

$$\dot{J}\left(\boldsymbol{\eta}_t;\boldsymbol{\eta}_0\right) = J\left(\boldsymbol{\eta}_t;\boldsymbol{\eta}_0\right)\sum_{i=1}^{N}\frac{\partial\,\dot{\eta}_{i,t}}{\partial\,\eta_{i,t}} \tag{2.94}$$

$$= J\left(\boldsymbol{\eta}_t;\boldsymbol{\eta}_0\right)\boldsymbol{\nabla}_{\boldsymbol{\eta}_t}\cdot\dot{\boldsymbol{\eta}}_t \tag{2.95}$$

A compressibility factor κ is introduced which describes the time dependency of the Jacobian. κ is a function of a phase space vector and the time.

$$\dot{J}\left(\boldsymbol{\eta}_t;\boldsymbol{\eta}_0\right) = J\left(\boldsymbol{\eta}_t;\boldsymbol{\eta}_0\right)\,\kappa\left(\boldsymbol{\eta}_t,t\right) \tag{2.96}$$

κ is defined as:

$$\kappa\left(\boldsymbol{\eta}_t,t\right) = \boldsymbol{\nabla}_{\boldsymbol{\eta}_t}\cdot\dot{\boldsymbol{\eta}}_t \tag{2.97}$$

$$= \frac{\dot{J}}{J} \tag{2.98}$$

$$= \frac{\mathrm{d}\ln J}{\mathrm{d}t} \tag{2.99}$$

The general solution of equation (2.96) is:

$$J\left(\boldsymbol{\eta}_t;\boldsymbol{\eta}_0\right) = \mathrm{e}^{\int_0^t \kappa(\boldsymbol{\eta}_s,s)\,\mathrm{d}s} \tag{2.100}$$

If $w\left(\boldsymbol{\eta}_t,t\right)$ is the primitive function of $\kappa\left(\boldsymbol{\eta}_t,t\right)$ then the Jacobian is given by:

$$J\left(\boldsymbol{\eta}_t;\boldsymbol{\eta}_0\right) = \mathrm{e}^{w(\boldsymbol{\eta}_t,t)-w(\boldsymbol{\eta}_0,0)} \tag{2.101}$$

Inserting the Jacobian into equation (2.83) leads to:

$$\mathrm{e}^{-w(\boldsymbol{\eta}_t,t)}\,\mathrm{d}\boldsymbol{\eta}_t = \mathrm{e}^{-w(\boldsymbol{\eta}_0,0)}\,\mathrm{d}\boldsymbol{\eta}_0 \tag{2.102}$$

Both sides of the equation are conserved by the dynamics, the phase space and an additional factor called the metric determinant factor $\sqrt{g(\boldsymbol{\eta}_t,t)}$. Written in this notation, the Jacobian is the ratio of two metric factors.

$$\sqrt{g(\boldsymbol{\eta}_t, t)}\,\mathrm{d}\boldsymbol{\eta}_t = \sqrt{g(\boldsymbol{\eta}_0, 0)}\,\mathrm{d}\boldsymbol{\eta}_0 \tag{2.103}$$

$$\sqrt{g(\boldsymbol{\eta}_t, t)} = \mathrm{e}^{-w(\boldsymbol{\eta}_t, t)} \tag{2.104}$$

$$\sqrt{g(\boldsymbol{\eta}_0, 0)} = \mathrm{e}^{-w(\boldsymbol{\eta}_0, 0)} \tag{2.105}$$

$$J(\boldsymbol{\eta}_t; \boldsymbol{\eta}_0) = \sqrt{\frac{g(\boldsymbol{\eta}_0, 0)}{g(\boldsymbol{\eta}_t, t)}} \tag{2.106}$$

If start and end time are equal then the ratio is one, equal to the Hamiltonian case. In some cases it is possible to obtain a flat phase space by a coordinate transformation. The metric factor would have the value one or is a constant.

With these relations generalized equations for the Liouville equation, ensemble averages and other properties can be constructed.

$$\frac{\partial\ (f\sqrt{g})}{\partial\, t} + \boldsymbol{\nabla}_{\boldsymbol{\eta}} \cdot (f\sqrt{g}\dot{\boldsymbol{\eta}}) = 0 \tag{2.107}$$

The generalized Liouville equation (2.107) is valid for equilibrium and non-equilibrium Hamiltonian or non-Hamiltonian systems. Furthermore it was shown that $\frac{\mathrm{d}f}{\mathrm{d}t} = 0$ is also valid for non-Hamiltonian systems and the metric factor itself is a solution of the generalized Liouville equation[67].

The generalized expression for ensemble averages (and similar for time averages) is:

$$< A >= \frac{\int \mathrm{d}\boldsymbol{\eta}\, \sqrt{g(\boldsymbol{\eta})}\, A(\boldsymbol{\eta})\, f(\boldsymbol{\eta})}{\int \mathrm{d}\boldsymbol{\eta}\, \sqrt{g(\boldsymbol{\eta})}\, f(\boldsymbol{\eta})} \tag{2.108}$$

The normal phase space distribution function (equation (2.69)) satisfies the generalized Liouville equation.

The microcanonical partition function, which is now invariant under coordinate transformations, is:

$$\Omega = \int \mathrm{d}\boldsymbol{\eta}\, \sqrt{g(\boldsymbol{\eta})}\, f(\boldsymbol{\eta}) \tag{2.109}$$

2.3.1. Summary

Relations known between Hamiltonian dynamic systems and classical statistical mechanics were extended to include non-Hamiltonian systems. A compressible phase space, with the compressibility $\kappa(\boldsymbol{\eta})$, was introduced. This means that the phase space has no predefined geometry. The non-Hamiltonian phase space is a general Riemannian manifold with a nontrivial metric. An invariant measure on the manifold can be calculated. If the phase space is incompressible, the extended equations will be equal to the original Hamiltonian equations. It was shown that the entropy of an equilibrium system is conserved in all cases[88].

The non-Hamiltonian theory was used in the Nosé-Hoover NVT section 5.1.2 on page 35, in the NPT section 5.2 on page 42 and in the GEMD section 6.2.4 on page 65.

3. Molecular Dynamics Integrators

To obtain a molecular dynamics trajectory, the equations of motions are solved numerically by the finite difference method. The system is defined by certain properties like the lattice positions, orientations, translational velocities, rotational velocities and intramolecular positions and velocities at an initial time step. A fundamental time step δt is given and the equations are solved in intervals of this time step. First, the solutions have to follow a classical trajectory for at least a correlation length. Second, it has to sample the phase space of the ensemble and to obey the conservation laws. Third, it should be time reversible. In general, a large time step with good conservation properties is desired.

The molecular dynamics method was first used to simulate hard spheres. The spheres move until they collide with another sphere, then the velocity vectors are corrected. This was done in 1957 by Alder and Wainwright to study phase transitions[2]. The first simulation using a Lennard-Jones potential was done by Rahman in 1964[66].

The time integrators of the Verlet family are the most popular in molecular dynamics. On the next pages an overview of the different algorithms and their advantages is given. The leap-frog algorithm is excluded because it is not used in this work and is rarely used for extended ensembles.

In the NVE ensemble, $\boldsymbol{q}$ is the position and $\dot{\boldsymbol{q}}$ the velocity of a particle. $\ddot{\boldsymbol{q}}$ is the acceleration of a particle given by $\boldsymbol{a}(t) = \frac{\boldsymbol{f}(t)}{m}$ where $\boldsymbol{f}$ is the force acting on the particle.

3.1. Verlet Integrator

Störmer published solutions to second order differential equations of the kind $\ddot{\boldsymbol{q}} = f(\boldsymbol{q})$ at the beginning of the last century[79]. He studied the trajectory of ions in the Van Allen belt. Verlet reinvented the algorithm and introduced it into the field of molecular dynamics[91]. Interesting is the fact that the same method was already used in Newton's *Principia* from 1687 to prove Kepler's second law. The algorithm has a long history and is used in different scientific fields. A geometrical analysis can be found in reference [29].

Today, the Verlet algorithm is in the literature also known as positional Verlet[1], Verlet-Störmer or Verlet-Størmer. Sometimes the leap-frog algorithm is also called Verlet.

[1] The counterpart to the velocity Verlet algorithm.

A forward and a backward Taylor expansion are the fundamental equations[3]:

$$\boldsymbol{q}(t+\delta t) = \boldsymbol{q}(t) + \dot{\boldsymbol{q}}(t)\,\delta t + \frac{1}{2}\,\ddot{\boldsymbol{q}}(t)\,\delta t^2 + \cdots \tag{3.1}$$

$$\boldsymbol{q}(t-\delta t) = \boldsymbol{q}(t) - \dot{\boldsymbol{q}}(t)\,\delta t + \frac{1}{2}\,\ddot{\boldsymbol{q}}(t)\,\delta t^2 - \cdots \tag{3.2}$$

By addition of both equations one obtains:

$$\boldsymbol{q}(t+\delta t) = 2\,\boldsymbol{q}(t) - \boldsymbol{q}(t-\delta t) + \ddot{\boldsymbol{q}}(t)\,\delta t^2 + \mathcal{O}\left(\delta t^4\right) \tag{3.3}$$

This is the Verlet algorithm. It is symmetric, time reversible and conserves the total energy and linear momentum.

The velocities are not part of equation (3.3). They are calculated by the following approximation:

$$\dot{\boldsymbol{q}} = \frac{\boldsymbol{q}(t+\delta t) - \boldsymbol{q}(t-\delta t)}{2\,\delta t} + \mathcal{O}\left(\delta t^2\right) \tag{3.4}$$

The initial values are $\boldsymbol{q}(t)$, $\boldsymbol{q}(t-\delta t)$ and $\ddot{\boldsymbol{q}}(t)$. Usually the positions and velocities at time t are known but not the previous positions. The previous positions are approximated by:

$$\boldsymbol{q}(t-\delta t) = \boldsymbol{q}(t) - \dot{\boldsymbol{q}}(t)\,\delta t \tag{3.5}$$

The Verlet algorithm is a symplectic numerical integrator (see section 2.1.2 on page 10). It was shown that this condition is important for the stability and accuracy of an integrator[30, 73]. Gear and Runge-Kutta algorithms are not symplectic and not time reversible[83].

If constraints are used in molecular systems, then the full step positions have to be corrected by an algorithm such as SHAKE[69].

3.2. Velocity Verlet Integrator

The velocity Verlet[80] algorithm is an equivalent version of the Verlet algorithm. The velocities are included in the integration equations. The authors write that their algorithm retains a

superior numerical precision. However, Batcho and Schlick[8] showed that the normal Verlet version is more stable if larger time steps are used. In general, all Verlet algorithms generate identical trajectories. But the different numerics leads to different rounding errors which depend on the numerical precision used.

The difference to the positional Verlet algorithm is that the positions are advanced in a single step and the velocities in two half steps.

$$\boldsymbol{q}(t+\delta t) = \boldsymbol{q}(t) + \dot{\boldsymbol{q}}(t)\,\delta t + \ddot{\boldsymbol{q}}(t)\,\frac{\delta t^2}{2} + \mathcal{O}\left(\delta t^4\right) \tag{3.6}$$

$$\dot{\boldsymbol{q}}\left(t+\frac{1}{2}\,\delta t\right) = \dot{\boldsymbol{q}}(t) + \ddot{\boldsymbol{q}}(t)\,\frac{\delta t}{2} + \mathcal{O}\left(\delta t^3\right) \tag{3.7}$$

$$\dot{\boldsymbol{q}}(t+\delta t) = \dot{\boldsymbol{q}}\left(t+\frac{1}{2}\,\delta t\right) + \ddot{\boldsymbol{q}}(t+\delta t)\,\frac{\delta t}{2} + \mathcal{O}\left(\delta t^3\right) \tag{3.8}$$

The initial values are $\boldsymbol{q}(t)$, $\dot{\boldsymbol{q}}(t)$ and $\ddot{\boldsymbol{q}}(t)$.

The advantage is that the algorithm is "self starting", with knowing the positions and velocities at the beginning, the iteration cycles can be started directly. If constraints are used, both the positions and velocities have to be corrected in an iterative way. For equation (3.6) and (3.7) a modified SHAKE[69] is used. To correct the velocities of equation (3.8), the so called RATTLE algorithm[5] is used.

3.3. General Molecular Dynamics Approximations

Several approximations and definitions are presumed in molecular dynamics simulations. The following conditions are used in the simulations shown later on.

System

Periodic boundary conditions are used to generate an infinite system without a surface.

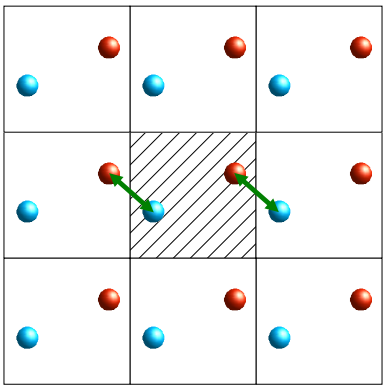


Figure 3.1.: Periodic Boundary Conditions in a Two Dimensional Lattice.

The minimum image convention is used to calculate pair interactions between nearest images.

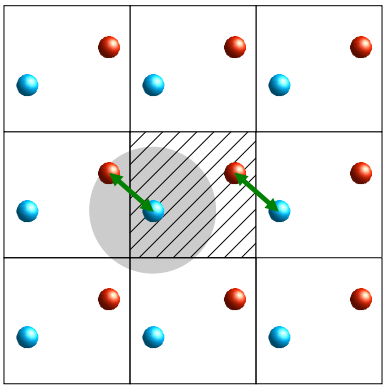


Figure 3.2.: Minimum Image Convention.

Potential

Usually the potential of a particle in the field of the surrounding particles is calculated using pair interactions. Three body and higher interactions are neglected. Pair additivity is assumed. The Born-Oppenheimer approximation is used. The potential is only a function of the distances of the nuclei.

To limit the interaction length between particles, a cut-off radius r_c is defined. If the distance is larger than this radius, the interaction is by definition zero.

$$\Phi(\boldsymbol{q}) = \begin{cases} \Phi(\boldsymbol{q}) & |\boldsymbol{q}_{ij}| \leq r_c \\ 0 & |\boldsymbol{q}_{ij}| > r_c \end{cases} \tag{3.9}$$

The cut-off radius must be in the range where the g-function decays to a uniform distribution.

In addition, the potential values are often shifted to get a zero potential value at the cut-off radius.

The error introduced by the cut-off radius is partially corrected by the long-range corrections, assuming an ideal mass distribution. Quantum corrections are sometimes used to correct the error due to the classical treatment.

Intramolecular Constraints

To simplify molecular systems, especially to get rid of fast intramolecular oscillations, angles and bond lengths are fixed.

4. Monte Carlo Method

In 1946, Stanislaw Ulam, a scientist of the Manhattan project in Los Alamos, was looking at the possibility to win the card game Solitaire. 52 cards are laid out in Canfield Solitaire. The combinatorial approach was too complicated, therefore he looked for a simpler method. A computational method using random numbers, the Monte Carlo method, was simple and efficient to solve this problem. The name Monte Carlo, was given by Metropolis because Ulam was interested in poker, and the capital of Monaco was a famous place for gambling. The method was later on used to calculate neutron diffusion processes[20].

Statistical sampling methods using random numbers or random processes have been used for centuries. In 1777 Buffon solved a geometrical probability problem, the needle problem. A needle with length l is dropped on a sheet of paper with parallel lines of distance d. The probability of hitting a line is $2\,l/\pi\,d$[3].

The advantage of Monte Carlo was that the generation of the random numbers and the sequential sampling could be efficiently done on the super computer MANIAC. In general, Monte Carlo can be used whenever a probability distribution is known and averaging over sample states is equal to the ensemble average. This is a similar condition to the ergodicity in molecular dynamics.

A simple example for the Monte Carlo method is the calculation of the lowest energy. An initial particle distribution is randomly moved in each cycle, if the new energy is lower, then the new configuration is accepted, otherwise it is dropped. The algorithm stops if the changes are small, indicating that a constant energy level is reached. Starting from different initial configurations and using different random number sequences help to find the global minimum.

In 1953 Metropolis *et al.*[54] presented a Monte Carlo method to sample a constant temperature ensemble. The algorithm is based on the minimal energy algorithm, in addition it allows configurations with higher energies by assigning a probability (equation (6.3) on page 54). If a uniform random number between zero and one is lower or equal to this probability, then the move is accepted. This is also called importance sampling.

The Metropolis approach can be used to sample any statistical ensemble given by a probability distribution. It is normally assumed that every configuration has the same probability. The sampling is a Markovian process, new configurations depend only on the previous one and all configurations are in one state of a finite state space. The Monte Carlo method gives no information about time dependent values and the magnitude of fluctuations.

5. Molecular Dynamics Ensembles

An ensemble is a collection of all microscopic systems at different states which have equal macroscopic thermodynamic properties. The fundamental ensemble used by molecular dynamics simulations is the microcanonical NVE ensemble. The number of particles, the volume and the energy are constant. The number of particles and the volume are constant by definition. The energy conservation is a consequence of using Newton's equations of motion.

A constant number of particles is usually wanted, but the restriction of a constant volume and energy are limiting factors. Experiments are often done in systems at constant temperature or/and pressure. The question is if and how one can get from the microcanonical to the canonical NVT or isothermal-isobaric NPT ensemble.

It was early known that Monte Carlo methods are able to sample different ensembles but it took a long time to be able to use these ensembles in molecular dynamics.

5.1. The Canonical Ensemble

In contrast to the microcanonical ensemble, the temperature is constant in the canonical ensemble whereas the energy is no longer a constant. Most experiments are done at constant temperature, using either an open system with a constant pressure in addition or a closed vessel surrounded by a heat bath. Some pressure values obtained by NVT simulations are shown in section 7.1.1 on page 75.

Nonetheless it is possible to sample quasi isothermal values in NVE simulations: during the equilibration the energy is changed to reach a constant temperature. The temperature drift in the production cycle is usually in the range from zero to a few Kelvin. By running several simulations, the target temperature can be interpolated. The quality of this method depends on the number of points and the fit function. A linear behavior of the pressure is often assumed. The disadvantage is that several runs have to be performed whereas in a real NVT ensemble one run, sampling canonical states, is sufficient.

At first pseudo isothermal ensembles were used. The velocities or momenta of the particles were scaled to get a constant temperature at each time step[97] according to equation (5.1) where g is the number of degrees of freedom. The problem of this technique is that the temperature fluctuations, known from statistical mechanics, are completely ignored. The rescaling has no physical meaning and the time reversibility is lost.

$$\frac{1}{2}\, g\, k_B\, T = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2\, m_i} \tag{5.1}$$

Another method, combining MD with Monte Carlo, was proposed by Andersen[4]. The velocities of the particles are changed by stochastic collisions. Similar problems known from the scaling procedure are introduced: the algorithm is no longer time reversible and the trajectories are not continuous in the phase space. Therefore it is unknown how different points of a trajectory are correlated and it is not possible to calculate transport properties.

In the same publication, Andersen introduced the piston method to control the pressure by an additional degree of freedom. However, Andersen was not able to enhance this extended Hamiltonian method to constant temperature dynamics.

The Berendsen thermostat[9] couples the system to an external heat bath. This generates correct average values but wrong temperature fluctuations. The method has no connection to statistical mechanics.

5.1.1. Nosé Dynamics

In 1984 Nosé published an important paper[57, 58] where he used Andersen's extended system method to construct a canonical ensemble. The additional degree of freedom is called s ($[s] = 1$), the effective "mass" of s is Q ($[Q] = \mathrm{kg\,m^2}$) and its momentum is p_s ($[p_s] = \mathrm{kg\,m^2\,s^{-1}}$).

Nosé transformed the real variables $\boldsymbol{q}_i'$, $\boldsymbol{p}_i'$ and t' into virtual variables

$$\boldsymbol{q}_i' = \boldsymbol{q}_i \tag{5.2}$$

$$\boldsymbol{p}_i' = \frac{\boldsymbol{p}_i}{s} \tag{5.3}$$

$$t' = \int^{t} \frac{\mathrm{d}t}{s} \tag{5.4}$$

$$\frac{\mathrm{d}\boldsymbol{q}_i'}{\mathrm{d}t'} = s\, \frac{\mathrm{d}\boldsymbol{q}_i}{\mathrm{d}t} \tag{5.5}$$

by a scaling of the time:

$$\mathrm{d}t' = \frac{\mathrm{d}t}{s} \tag{5.6}$$

The extended Hamiltonian was postulated as:

$$\mathcal{H}_{\text{Nosé}} = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2\, m_i\, s^2} + \Phi(\boldsymbol{q}) + \frac{p_s^2}{2\, Q} + g\, k_B\, T_E \ln s \tag{5.7}$$

The equations of motion are obtained by the usual Hamiltonian formalism (see equations (2.40) and (2.41)).

$$\begin{aligned}
\dot{\boldsymbol{q}}_i &= \frac{\partial\, \mathcal{H}_{\text{Nosé}}}{\partial\, \boldsymbol{p}_i} \\
&= \frac{\boldsymbol{p}_i}{m_i\, s^2} && (5.8) \\
\dot{\boldsymbol{p}}_i &= -\frac{\partial\, \mathcal{H}_{\text{Nosé}}}{\partial\, \boldsymbol{q}_i} \\
&= -\frac{\partial\, \Phi(\boldsymbol{q})}{\partial\, \boldsymbol{q}_i} && (5.9) \\
\dot{s} &= \frac{\partial\, \mathcal{H}_{\text{Nosé}}}{\partial\, p_s} \\
&= \frac{p_s}{Q} && (5.10) \\
\dot{p}_s &= -\frac{\partial\, \mathcal{H}_{\text{Nosé}}}{\partial\, s} \\
&= \frac{\sum\limits_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m_i\, s^2} - g\, k_B\, T_E}{s} && (5.11)
\end{aligned}$$

The same conservation laws are valid and the time derivative of the Hamiltonian is zero. The extended Hamiltonian (5.7) produces a microcanonical ensemble. Therefore the partition function Ω of the entire ensemble in virtual variables is (see equation (2.75))

$$\Omega_{\text{Nosé}} = \int \mathrm{d}p_s \int \mathrm{d}s \int \mathrm{d}^N \boldsymbol{p} \int \mathrm{d}^N \boldsymbol{q}\, \delta \left[\mathcal{H}_0 \left(\frac{\boldsymbol{p}}{s}, \boldsymbol{q} \right) + \frac{p_s^2}{2\, Q} + g\, k_B\, T_E \ln s - E \right] \tag{5.12}$$

where the prefactors were omitted.

The coordinates and momenta of the real particles are transformed to real variables. The following partition function is obtained:

$$\Omega_{\text{Nosé}} = \int \mathrm{d}p_s \int \mathrm{d}s\, s^{3\,N} \int \mathrm{d}^N \boldsymbol{p}' \int \mathrm{d}^N \boldsymbol{q}'\, \delta \left[\mathcal{H}_0(\boldsymbol{p}', \boldsymbol{q}') + \frac{p_s^2}{2\, Q} + g\, k_B\, T_E \ln s - E \right] \tag{5.13}$$

The δ function is zero for only one certain value of s, therefore the equivalence relation[51]

$$\delta\left[f(s)\right] = \frac{\delta(s - s_0)}{f'(s_0)} \tag{5.14}$$

can be used, where s_0 is the zero of $f(s)$.

$$\begin{aligned}
s_0 &= \exp\left(-\frac{\mathcal{H}_0(\boldsymbol{p}', \boldsymbol{q}') + \frac{p_s^2}{2\,Q} - E}{g\,k_B\,T_E}\right) && (5.15)\\
f(s) &= \mathcal{H}_0(\boldsymbol{p}', \boldsymbol{q}') + \frac{p_s^2}{2\,Q} + g\,k_B\,T_E \ln s - E && (5.16)\\
f'(s) &= \frac{g\,k_B\,T_E}{s} && (5.17)\\
\Omega_{\text{Nosé}} &= \frac{1}{g\,k_B\,T_E} \int \mathrm{d}p_s \int \mathrm{d}^N\boldsymbol{p}' \int \mathrm{d}^N\boldsymbol{q}' \int \mathrm{d}s\, s^{3\,N+1}\,\delta\left(s - s_0\right) && (5.18)
\end{aligned}$$

The s integral is solved using the delta function property

$$\begin{aligned}
\int \mathrm{d}s\, f(s)\,\delta(s - s_0) &= f(s_0) && (5.19)\\
\int \mathrm{d}s\, s^{3\,N+1}\,\delta(s - s_0) &= s_0^{3\,N+1} && (5.20)
\end{aligned}$$

leading to:

$$\begin{aligned}
\Omega_{\text{Nosé}} &= \frac{1}{g\,k_B\,T_E} \exp\left(\frac{3\,N+1}{g}\frac{E}{k_B\,T_E}\right) \int \mathrm{d}p_s \exp\left(-\frac{3\,N+1}{g}\frac{p_s^2}{2\,Q\,k_B\,T_E}\right) \int \mathrm{d}^N\boldsymbol{p}' \\
&\quad \times \int \mathrm{d}^N\boldsymbol{q}' \exp\left(-\frac{3\,N+1}{g}\frac{\mathcal{H}_0(\boldsymbol{p}', \boldsymbol{q}')}{k_B\,T_E}\right) && (5.21)
\end{aligned}$$

If g is set to $3\,N + 1$, then the partition function is reduced to the partition function of a canonical ensemble (equation (2.78)) plus a constant factor:

$$\Omega_{\text{Nosé}} = C \int \mathrm{d}^N\boldsymbol{p}' \int \mathrm{d}^N\boldsymbol{q}' \exp\left(-\frac{\mathcal{H}_0(\boldsymbol{p}', \boldsymbol{q}')}{k_B\,T_E}\right) \tag{5.22}$$

The problem of virtual time sampling is that real time intervals are unequal. Therefore Nosé transformed the equations of motion to real variables (the derivation is not shown here). The value of g in real time has to be $3\,N$. The obtained energy conservation equation is no longer a Hamiltonian and the equations of motion cannot be obtained by derivatives of the Hamiltonian.

$$\dot{\boldsymbol{q}}'_i = \frac{\boldsymbol{p}'_i}{m_i} \tag{5.23}$$

$$\dot{\boldsymbol{p}}'_i = -\frac{\partial\,\Phi(\boldsymbol{q}')}{\partial\,\boldsymbol{q}'_i} - \frac{s'\,p'_s\,\boldsymbol{p}'_i}{Q} \tag{5.24}$$

$$\dot{s}' = \frac{s'^2\,p'_s}{Q} \tag{5.25}$$

$$\dot{p}'_s = \frac{\sum\limits_{i=1}^{N} \frac{\boldsymbol{p}'^2_i}{m_i} - g\,k_B\,T_E}{s'} - \frac{s'\,p_s^2}{Q} \tag{5.26}$$

$$\mathcal{H}'_{\text{Nosé}} = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2\,m_i} + \Phi(\boldsymbol{q}') + \frac{s'^2\,p'^2_s}{2\,Q} + g\,k_B\,T_E \ln s \tag{5.27}$$

Where the real variables of the extended parameters are defined as:

$$s' = s \tag{5.28}$$

$$p'_s = \frac{p_s}{s} \tag{5.29}$$

Periodic Boundary Conditions

If periodic boundary conditions are used, then the degrees of freedom used to calculate the temperature are decreased by D (the spatial dimension). This is due to the momentum conservation law. In real time, g would have the value $D\,(N-1)$. Important is the fact that the canonical ensemble is only generated if the total momentum is always zero[14]. The total momentum is a conserved quantity of the equations of motion, therefore the initial total momentum has to be zero. Consequences of a non-zero center of mass momentum are given in reference [32].

In SimMol the momenta are corrected in the input program by adjusting the velocity vectors:

$$\boldsymbol{P} = \sum_{i=1}^{N} m_i\,\boldsymbol{v}_i \tag{5.30}$$

$$M = \sum_{i=1}^{N} m_i \tag{5.31}$$

$$\boldsymbol{v}_i' = \boldsymbol{v}_i - \frac{\boldsymbol{P}}{M} \tag{5.32}$$

The generated ensemble is no longer an NVT ensemble but a (N-1)VT ensemble. The decrease by one in the particle number is irrelevant if thermodynamic averages are calculated, but changes the dynamic properties and the relaxation time of the system. If the system is large enough, the differences to the NVT ensemble will vanish. For small cluster systems this has to be considered.

The instantaneous temperature has to be calculated using the decreased degrees of freedom, otherwise the average temperature will not be equal to the external temperature.

$$T(t) = \frac{2}{D\,(N-1)\,k_B}\,\mathcal{K}(t) \tag{5.33}$$

$$\mathcal{K}(t) = \sum_{i=1}^{N} \frac{1}{2}\,m_i\,\boldsymbol{v}_i^2(t) \tag{5.34}$$

The zero total momentum condition and no external forces are special cases. To generate the canonical ensemble for general systems, several extensions have been published. The Nosé-Hoover chain method adds a chain of connected thermostats[49]. The same approach can be used for the barostat in the NPT ensemble. The algorithm is ideal for small system, as used in Car-Parrinello[13] simulations, because it optimizes the ergodicity. Other algorithms[86] like the Melchionna-Ciccoti-Holian algorithm include the center of mass momentum in the equations of motion. The Martyna-Tobias-Klein algorithm is capable of building an NPT ensemble if the total momentum is not zero. In general, all algorithms generate the (N-1)VT or (N-1)PT ensemble if no external forces are present and the total linear momentum is zero.

A further consequence of using periodic boundary conditions is that the total angular momentum is no longer a conserved quantity.

A hypothetical problem of the Nosé approach is that a system at zero Kelvin cannot be thermostated to another temperature. This can easily be avoided by adjusting the initial velocities, to start from a non-zero temperature.

5.1.2. Nosé-Hoover Dynamics

Hoover transformed Nosé's method to a more general formulation without scaling the time[36]. In fact he introduced the non-Hamiltonian scheme to molecular dynamics which is the most common method today. In this context non-Hamiltonian means that the equations of motion

are not derived from a Hamiltonian (see section 2.3 on page 16). Hoover's method is called Nosé-Hoover in the literature[21].

Hoover started with Nosé's equations of motion (5.8)-(5.11) in virtual time. The degrees of freedom g in equation (5.11) are written as $g = X + 1$. X are the $3\,N$ degrees of freedom of the real particles. These equations are simplified by unscaling the time with $\mathrm{d}t_{new} \equiv \mathrm{d}t_{old}/s$ to:

$$\dot{\boldsymbol{q}}_i = \frac{\boldsymbol{p}_i}{m_i\, s} \tag{5.35}$$

$$\dot{\boldsymbol{p}}_i = -s\,\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}_i} \tag{5.36}$$

$$\dot{s} = \frac{s\,p_s}{Q} \tag{5.37}$$

$$\dot{p}_s = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m_i\, s^2} - (X+1)\,k_B\,T_E \tag{5.38}$$

To eliminate the variable s, the time derivative of equation (5.35) is calculated and combined with the other equations of motion

$$\ddot{\boldsymbol{q}}_i = \frac{\dot{\boldsymbol{p}}_i}{m_i\, s} - \frac{\boldsymbol{p}_i}{m_i\, s}\frac{\dot{s}}{s} \tag{5.39}$$

$$= -\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}_i}\,\frac{1}{m_i} - \frac{\dot{\boldsymbol{q}}_i\,p_s}{Q} \tag{5.40}$$

$$\equiv -\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}_i}\,\frac{1}{m_i} - \xi\dot{\boldsymbol{q}}_i \tag{5.41}$$

where the thermodynamic friction coefficient ξ ($[\xi] = \mathrm{s}^{-1}$) was introduced. Its value and time derivative are:

$$\xi = \frac{p_s}{Q} \tag{5.42}$$

$$\dot{\xi} = \frac{\sum\limits_{i=1}^{N} m_i\,\dot{\boldsymbol{q}}_i^2 - (X+1)\,k_B\,T_E}{Q} \tag{5.43}$$

In order to get canonical equations of motion, the momenta are redefined as $\boldsymbol{p}_i \equiv m_i\,\dot{\boldsymbol{q}}_i$ and $X+1$ is replaced by X.

$$\dot{\boldsymbol{q}}_i = \frac{\boldsymbol{p}_i}{m_i} \tag{5.44}$$

$$\dot{\boldsymbol{p}}_i = -\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}_i} - \xi\,\boldsymbol{p}_i \tag{5.45}$$

$$\dot{\xi} = \frac{\sum\limits_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m_i} - X\,k_B\,T_E}{Q} \tag{5.46}$$

These are Hoover's equations of motion. The variable s was eliminated and all time intervals are of equal size. The friction coefficient ξ, a rate or velocity, scales the forces to reach an average kinetic energy which is equal to the external temperature.

Conventional Analysis

This section is called "conventional analysis", this is how it was done by Hoover in the original publication. He analyzed not the complete phase space but a subset using a postulated partition function. The method is not wrong but leads to incomplete results. The complete phase space behavior is shown in the non-Hamiltonian part.

The equations of motion must satisfy Liouville's equation (equation (2.66) on page 12) which is valid for a Hamiltonian and in general is based on the continuity equation (2.65). The probability flow must be independent of time.

$$\frac{\partial\,f}{\partial\,t} + \dot{\boldsymbol{q}}\,\frac{\partial\,f}{\partial\,\boldsymbol{q}} + \dot{\boldsymbol{p}}\,\frac{\partial\,f}{\partial\,\boldsymbol{p}} + \dot{\xi}\,\frac{\partial\,f}{\partial\,\xi} + f\left(\frac{\partial\,\dot{\boldsymbol{q}}}{\partial\,\boldsymbol{q}} + \frac{\partial\,\dot{\boldsymbol{p}}}{\partial\,\boldsymbol{p}} + \frac{\partial\,\dot{\xi}}{\partial\,\xi}\right) = 0 \tag{5.47}$$

The individual terms in equation (5.47) using a probability density function $f_{\mathrm{NVT}}(\boldsymbol{q},\boldsymbol{p},\xi)$ of the form

$$f_{\mathrm{NVT}} \propto \exp\left(-\frac{\Phi(\boldsymbol{q}) + \sum\limits_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2\,m_i} + \frac{1}{2}\,Q\,\xi^2}{k_B\,T_E}\right) \tag{5.48}$$

lead to the following expressions:

$$\frac{\partial\,f}{\partial\,t} = 0 \tag{5.49}$$

$$\dot{\boldsymbol{q}}\,\frac{\partial\,f}{\partial\,\boldsymbol{q}} = \frac{f}{k_B\,T_E}\sum_{i=1}^{N}\frac{\boldsymbol{p}_i}{m_i}\,\boldsymbol{f}_i(\boldsymbol{q}) \tag{5.50}$$

$$\dot{\boldsymbol{p}}\frac{\partial f}{\partial \boldsymbol{p}} = \frac{f}{k_B T_E}\sum_{i=1}^{N}(-\boldsymbol{f}_i(\boldsymbol{q}) + \xi\,\boldsymbol{p}_i)\frac{\boldsymbol{p}_i}{m_i} \tag{5.51}$$

$$\dot{\xi}\frac{\partial f}{\partial \xi} = \frac{f}{k_B T_E}\left(-\sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{m_i} + X\,k_B T_E\right)\xi \tag{5.52}$$

$$\frac{\partial \dot{\boldsymbol{q}}}{\partial \boldsymbol{q}} = 0 \tag{5.53}$$

$$\frac{\partial \dot{\boldsymbol{p}}}{\partial \boldsymbol{p}} = -D\sum_{i=1}^{N}\xi \tag{5.54}$$

$$= -X\,\xi \tag{5.55}$$

$$\frac{\partial \dot{\xi}}{\partial \xi} = 0 \tag{5.56}$$

To fulfill Liouville's equation, the constant X has to be equal to the number of independent degrees of freedom.

$$X = D\,N \tag{5.57}$$

Assuming periodic boundary conditions, the value of X has to be $D\,(N-1)$. Again, as in the Nosé algorithm, the total momentum has to be zero to generate a (N-1)VT ensemble.

Non-Hamiltonian Analysis

The conserved quantity, a pseudo Hamiltonian, is[86]:

$$\mathcal{H}' = \sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2\,m_i} + \Phi(\boldsymbol{q}) + \frac{1}{2}\,Q\,\xi^2 + X\,k_B T_E\,\eta \tag{5.58}$$

$$= H_0(\boldsymbol{p},\boldsymbol{q}) + \frac{1}{2}\,Q\,\xi^2 + X\,k_B T_E\,\eta \tag{5.59}$$

Here η ($[\eta] = 1$) is the coordinate of the thermostat ($\dot{\eta} = \xi$). Its value shows how much energy was already transferred from or to the heat bath. ξ is the "velocity" assigned to the thermostat mass. The η term is the potential energy and the ξ term the kinetic energy of the additional degree of freedom belonging to the thermostat. The conserved quantity is related to the Helmholtz free energy given by equation (2.79).

The phase space used in Hoover's Liouville equation is not complete, because the parameter η is missing, which is part of the Hamiltonian. As shown in the non-Hamiltonian section 2.3, the

phase space is compressible and the generalized Liouville equation (equation (2.107) on page 19) has to be used. The compressibility of the $2\,D\,N+2$ dimensional phase space is[86]:

$$\begin{aligned}
\kappa(\boldsymbol{\eta}) &= \boldsymbol{\nabla}_{\boldsymbol{\eta}} \cdot \dot{\boldsymbol{\eta}} && (5.60)\\
&= \sum_{i=1}^{2DN+2} \boldsymbol{\nabla}_{\boldsymbol{\eta}_i} \cdot \dot{\boldsymbol{\eta}}_i && (5.61)\\
\frac{\partial\,\dot{\boldsymbol{q}}}{\partial\,\boldsymbol{q}} &= 0 && (5.62)\\
\frac{\partial\,\dot{\boldsymbol{p}}}{\partial\,\boldsymbol{p}} &= -\xi\,D\,N && (5.63)\\
\frac{\partial\,\dot{\xi}}{\partial\,\xi} &= 0 && (5.64)\\
\frac{\partial\,\dot{\eta}}{\partial\,\eta} &= \frac{\partial\,\xi}{\partial\,\eta} = 0 && (5.65)\\
\kappa(\boldsymbol{\eta}) &\equiv -\xi\,D\,N && (5.66)
\end{aligned}$$

It follows that the metric determinant factor $\sqrt{g}$ is:

$$\begin{aligned}
\kappa(\boldsymbol{\eta}) &= \dot{w}(\boldsymbol{\eta}) && (5.67)\\
\dot{w}(\boldsymbol{\eta}) &= -\xi\,D\,N && (5.68)\\
w(\boldsymbol{\eta}) &= -\eta\,D\,N && (5.69)\\
\sqrt{g(\boldsymbol{\eta})} &= \mathrm{e}^{-w(\boldsymbol{\eta})} && (5.70)\\
&\equiv \mathrm{e}^{\eta\,D\,N} && (5.71)
\end{aligned}$$

If an effective coupling between the system and the thermostat exists, ξ fluctuates around zero with an average value of zero. Due to the oscillation, the highest probabilities are at the inflection points. At the equilibrium temperature, the metric determinant factor oscillates around a constant value.

Knowing the metric determinant factor, the generalized partition function is calculated:

$$\begin{aligned}
\Omega_{\text{Nosé}-\text{Hoover}} &= \int \mathrm{d}^N\boldsymbol{p} \int\limits_{D(V)} \mathrm{d}^N\boldsymbol{q} \int \mathrm{d}\xi \int \mathrm{d}\eta\,\mathrm{e}^{\eta\,D\,N}\\
&\quad \times\delta\left[\mathcal{H}_0(\boldsymbol{p},\boldsymbol{q}) + \frac{1}{2}\,Q\,\xi^2 + X\,k_B\,T_E\,\eta - E\right] && (5.72)
\end{aligned}$$

Here $D(V)$ is the spatial domain given by the volume.

The integration over η using Dirac's delta function is done in the same way as for Nosé's partition function.

$$\eta_0 = \frac{E - \mathcal{H}_0(\boldsymbol{p},\boldsymbol{q}) - \frac{1}{2}\,Q\,\xi^2}{X\,k_B\,T_E} \tag{5.73}$$

$$f'(\eta) = X\,k_B\,T_E \tag{5.74}$$

$$\begin{aligned}\Omega_{\text{Nosé}-\text{Hoover}} = {} & \frac{1}{X\,k_B\,T_E}\,\exp\left(\frac{D\,N}{X}\,\beta\,E\right) \\ & \times \int \mathrm{d}^N\boldsymbol{p} \int\limits_{D(V)} \mathrm{d}^N\boldsymbol{q}\,\exp\left(-\frac{D\,N}{X}\,\beta\,\mathcal{H}_0(\boldsymbol{p},\boldsymbol{q})\right) \\ & \times \int \mathrm{d}\xi\,\exp\left(-\frac{D\,N}{X}\,\frac{\beta}{2}\,Q\,\xi^2\right)\end{aligned} \tag{5.75}$$

The partition function is proportional to the canonical partition function (equation (5.76)), if X is equal to $D\,N$.

$$\Omega_{\text{NVT}} = \frac{C_N}{h^{D\,N}} \int \mathrm{d}^N\boldsymbol{p} \int\limits_{D(V)} \mathrm{d}^N\boldsymbol{q}\,\exp\left(-\beta\,\mathcal{H}(\boldsymbol{p},\boldsymbol{q})\right) \tag{5.76}$$

Where C_N is a combinatorial factor.

The Nosé-Hoover equations of motion generate the canonical distribution in the physical subsystem, if the Hamiltonian (5.58) is the only conservation law.

5.1.3. Choice of Q

The coupling of the thermostat to the particles is determined by the mass parameter Q. In the non-Hamiltonian part above it was shown that the isothermal partition function is generated for any value of Q. If a time sampling is done, as in molecular dynamics, the total length of the simulation and the quality of the expectation values depends on the value of Q. An optimal range for Q does exist. If Q is too large, then the energy exchange between the heat bath and the degrees of freedom is too small. The thermalization is inefficient or negligible. On the other hand, a too small Q leads to a high frequency coupling to the heat bath. In extreme cases, the time step has to be decreased to avoid resonance effects. Due to the high frequency, the sampled phase space is mostly not equal to the canonical one and the sampling is again inefficient. Therefore it is important to choose an appropriate value for Q to get a canonical distribution[35]. One rule of thumb is to select a rate which is equal to a characteristic vibrational frequency.

In dense fluids the trajectory of particles often changes and the motion can be considered as an oscillation of the particle in the surrounding field. The thermostat has then the same frequency as the temperature fluctuation of the system. A too strong coupling would just rescale the momenta without giving the system the time to evolve and adapt to the new temperature. By using a too weak coupling, the system fluctuations dominate the instantaneous temperature and the difference between the average and external temperature would increase.

The temperature fluctuation, which is larger than in the microcanonical ensemble, is in first approximation a harmonic oscillation. The period of the fluctuation $\tau(T)$ is given by the following equation[50]

$$\tau(T) = 2\,\pi\sqrt{\frac{Q}{X\,k_B\,T_E}} \tag{5.77}$$

with the frequency:

$$f = \frac{1}{\tau(T)} \tag{5.78}$$

$$= \frac{1}{2\,\pi}\sqrt{\frac{X\,k_B\,T_E}{Q}} \tag{5.79}$$

The thermostat's frequency is a function of the degrees of freedom. Changing the system size needs an adjustment of the mass Q to use the same frequency:

$$Q_2 = Q_1\,\frac{X_2}{X_1} \tag{5.80}$$

It is now obvious why the parameter Q is called a pseudo mass: formally Q is a replacement of the mass parameter m of a harmonic spring.

$$\tau = 2\,\pi\sqrt{\frac{m}{k}} \tag{5.81}$$

5.2. The Isothermal-Isobaric Ensemble

Andersen[4] introduced the isobaric method to molecular dynamics by adding an additional degree of freedom which scaled the coordinates and therefore the volume. The basic principle

is that the volume is inversely proportional to the pressure. The volume oscillates to obtain an average pressure which is equal to the external pressure. The density and energy are not constant in this ensemble. Ciccotti *et al.* conclude that internal metastabilities, due to the finite size of the simulation cell, are removed through the presence of volume fluctuations[15]. Nosé[57, 58] added Andersen's barostat to his canonical ensemble to get the isothermal-isobaric ensemble. Hoover extended his equations of motion (5.44)-(5.46) to obtain the same ensemble[36, 37].

The methods discussed here are only valid for homogeneous fluids, *e.g.* liquids or gases. If inhomogeneous systems have to be simulated then an anisotropic scaling of the simulation box has to be used with flexible lengths and angles[64, 50].

Hoover introduced reduced coordinates $\boldsymbol{r}_i$ (called $\boldsymbol{x}$ by Hoover) which are independent of the simulation box length and the volume. D is the number of spatial dimensions in the system and V the instantaneous volume.

$$\boldsymbol{r}_i \equiv \frac{\boldsymbol{q}_i}{V^{\frac{1}{D}}} \tag{5.82}$$

It is important that these coordinates are in the range $0 \leq r_i < 1$ so that all particles are in the unit cell.

Hoover used a relaxation time τ in his publication. Here a consistent formulation will be introduced to describe the barostat's degree of freedom[40] and which is later extended for the Gibbs ensemble method. Instead of the relaxation time, a second "mass" Q_P is introduced with the same units as the thermostat mass Q. Constant prefactors are incorporated into Q_P. All barostat variables have a P in the subscript. They are related to Hoover's variables in the following way:

Hoover	**SimMol**
$\dot{\epsilon}$	ξ_P
$\ddot{\epsilon}$	$\dot{\xi}_P$
τ	$\sqrt{\frac{Q_P}{D\,k_B\,T_E}}$

The equations of motion are:

$$\dot{\boldsymbol{r}}_i = \frac{\boldsymbol{p}_i}{m_i\,V^{\frac{1}{D}}} \tag{5.83}$$

$$\dot{\boldsymbol{p}}_i = -\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}_i} - (\xi_P + \xi)\,\boldsymbol{p}_i \tag{5.84}$$

$$\dot{\xi} = \frac{\sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m_i} - X\,k_B\,T_E}{Q} \tag{5.85}$$

$$\dot{V} = D\,V\,\xi_P \tag{5.86}$$

$$\dot{\xi}_P = \frac{D\,V(P - P_E)}{Q_P} \tag{5.87}$$

ξ_P ($[\xi_P] = \mathrm{s}^{-1}$) is the velocity and $\dot{\xi}_P$ the acceleration of the barostat's degree of freedom. Andersen and Nosé used $\dot{V}$ for the equations of motion, Hoover used the strain rate ξ_P. Both build qualitatively the same ensemble with different equations of motion.

The strain rate ξ_P, the velocity of the volume change, acts on the force vectors similar to the friction coefficient. The volume oscillates to obtain an average instantaneous pressure which is equal to the external pressure.

By selecting a large mass Q_P, the influence of the barostat on the particles is negligibly small and the normal Nosé-Hoover NVT ensemble is generated. The same can be done with the thermostat mass Q, whereby a NPE ensemble is constructed (Andersen's ensemble). By decoupling both masses, one obtains the usual NVE ensemble.

Conventional Analysis

The equations of motion must again satisfy Liouville's theorem and therefore the continuity equation (2.65):

$$\begin{aligned}\frac{\partial f}{\partial t} + \dot{\boldsymbol{r}}\,\frac{\partial f}{\partial \boldsymbol{r}} + \dot{\boldsymbol{p}}\,\frac{\partial f}{\partial \boldsymbol{p}} + \dot{\xi}\,\frac{\partial f}{\partial \xi} + \dot{V}\,\frac{\partial f}{\partial V} + \dot{\xi}_P\,\frac{\partial f}{\partial \xi_P} \\ + f\left(\frac{\partial \dot{\boldsymbol{r}}}{\partial \boldsymbol{r}} + \frac{\partial \dot{\boldsymbol{p}}}{\partial \boldsymbol{p}} + \frac{\partial \dot{\xi}}{\partial \xi} + \frac{\partial \dot{V}}{\partial V} + \frac{\partial \dot{\xi}_P}{\partial \xi_P}\right) &= 0\end{aligned} \tag{5.88}$$

The proposed probability density function $f_{\mathrm{NPT}}(\boldsymbol{r}, \boldsymbol{p}, \xi, V, \xi_P)$ is proportional to:

$$f_{\mathrm{NPT}} \propto V^{N-1} \exp\left(-\frac{\Phi\left(\boldsymbol{r}\,V^{\frac{1}{D}}\right) + \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2\,m_i} + \frac{1}{2}\,Q\,\xi^2 + P_E\,V + \frac{1}{2}\,Q_P\,\xi_P^2}{k_B\,T_E}\right) \tag{5.89}$$

The individual terms are:

$$\frac{\partial f}{\partial t} = 0 \tag{5.90}$$

$$\dot{\boldsymbol{r}}\,\frac{\partial f}{\partial \boldsymbol{r}} = \frac{f}{k_B\,T_E}\sum_{i=1}^{N}\frac{\boldsymbol{p}_i}{m_i}\,\boldsymbol{f}_i(\boldsymbol{q}) \tag{5.91}$$

$$\dot{\boldsymbol{p}}\,\frac{\partial f}{\partial \boldsymbol{p}} = \frac{f}{k_B\,T_E}\sum_{i=1}^{N}(-\boldsymbol{f}_i(\boldsymbol{q}) + (\xi+\xi_P)\,\boldsymbol{p}_i)\,\frac{\boldsymbol{p}_i}{m_i} \tag{5.92}$$

$$\dot{\xi}\,\frac{\partial f}{\partial \xi} = \frac{f}{k_B\,T_E}\left(-\sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{m_i} + X\,k_B\,T_E\right)\xi \tag{5.93}$$

$$\dot{\xi}_P\,\frac{\partial f}{\partial \xi_P} = -\frac{f}{k_B\,T_E}\,D\,V(P-P_E)\,\xi_P \tag{5.94}$$

$$\dot{V}\,\frac{\partial f}{\partial V} = \frac{f}{k_B\,T_E}\,D\,V\,\xi_P\left(\frac{(N-1)\,k_B\,T_E}{V} - P_E + \frac{1}{D\,V}\,\boldsymbol{q}\cdot\boldsymbol{f}(\boldsymbol{q})\right) \tag{5.95}$$

$$\frac{\partial\,\dot{\boldsymbol{r}}}{\partial\,\boldsymbol{r}} = 0 \tag{5.96}$$

$$\frac{\partial\,\dot{\boldsymbol{p}}}{\partial\,\boldsymbol{p}} = -D\sum_{i=1}^{N}(\xi+\xi_P) \tag{5.97}$$

$$= -D\,N\,(\xi+\xi_P) \tag{5.98}$$

$$\frac{\partial\,\dot{\xi}}{\partial\,\xi} = 0 \tag{5.99}$$

$$\frac{\partial\,\dot{V}}{\partial\,V} = D\,\xi_P \tag{5.100}$$

$$\frac{\partial\,\dot{\xi}_P}{\partial\,\xi_P} = 0 \tag{5.101}$$

For the partial derivative of the potential with respect to the volume (equation (5.106)) the following relations were used. All used vectors are collinear, therefore a vector division is possible, which is equal to the division of their length.

$$\boldsymbol{q} = V^{\frac{1}{D}}\,\boldsymbol{r} \tag{5.102}$$

$$V = \left(\frac{\boldsymbol{q}}{\boldsymbol{r}}\right)^{D} \tag{5.103}$$

$$\frac{\partial V}{\partial\,\boldsymbol{q}} = \frac{D\,\boldsymbol{q}^{D-1}}{\boldsymbol{r}^{D}} \tag{5.104}$$

$$= \frac{D\,V}{\boldsymbol{q}} \tag{5.105}$$

$$\frac{\partial\,(\Phi(\boldsymbol{q}))}{\partial\,V} = \frac{\partial\,\boldsymbol{q}}{\partial\,V}\cdot\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}} \tag{5.106}$$

$$= \frac{\boldsymbol{q}}{D\,V}\cdot\frac{\partial\,\Phi(\boldsymbol{q})}{\partial\,\boldsymbol{q}} \tag{5.107}$$

$$= -\frac{\boldsymbol{q}}{D\,V}\cdot\boldsymbol{f}(\boldsymbol{q}) \tag{5.108}$$

$$= -\frac{\mathcal{W}}{V} \tag{5.109}$$

Where the virial $\mathcal{W}$ was introduced.

With $X = D\,N$, the non-vanishing terms of Liouville's equation are the definition of the pressure, the virial equation divided by V.

$$P = \frac{\sum\limits_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m_i} + \boldsymbol{q} \cdot \boldsymbol{f}(\boldsymbol{q})}{D\,V} \tag{5.110}$$

The Nosé-Hoover equations of motion satisfy the probability conservation law. This condition is not sufficient to generate the isothermal-isobaric ensemble.

Non-Hamiltonian Analysis

It is the first time that the non-Hamiltonian formalism is used to prove that Hoover's equations of motion generate the proposed probability density function and that a quasi isothermal-isobaric partition function is obtained.

The conserved energy, given by a pseudo Hamiltonian, is:

$$\begin{aligned} \mathcal{H}' &= \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2\,m_i} + \Phi(\boldsymbol{r}\,V^{\frac{1}{D}}) + \frac{1}{2}\,Q\,\xi^2 + X\,k_B\,T_E\,\eta + \frac{1}{2}\,Q_P\,\xi_P^2 + P_E\,V & (5.111)\\ &= \mathcal{H}_0(\boldsymbol{p}, \boldsymbol{r}\,V^{\frac{1}{D}}) + \frac{1}{2}\,Q\,\xi^2 + X\,k_B\,T_E\,\eta + \frac{1}{2}\,Q_P\,\xi_P^2 + P_E\,V & (5.112) \end{aligned}$$

The conserved quantity is related to the Gibbs free energy, given by equation (2.82).

The isothermal-isobaric equations of motion are again analyzed with the non-Hamiltonian theory (see section 2.3 on page 16). The phase space has the dimension $2\,D\,N + 4$. The barostat degree of freedom adds two additional variables to the phase space, the volume V is part of the potential energy term and ξ_P of the kinetic energy term in the Hamiltonian (5.112).

The compressibility of the phase space $\kappa\,(\boldsymbol{\eta})$ is:

$$\begin{aligned} \kappa(\boldsymbol{\eta}) &= \boldsymbol{\nabla}_{\boldsymbol{\eta}} \cdot \dot{\boldsymbol{\eta}} & (5.113)\\ &= \sum_{i=1}^{2DN+4} \boldsymbol{\nabla}_{\boldsymbol{\eta}_i} \cdot \dot{\boldsymbol{\eta}}_i & (5.114)\\ \frac{\partial\,\dot{\boldsymbol{r}}}{\partial\,\boldsymbol{r}} &= 0 & (5.115) \end{aligned}$$

$$\frac{\partial \dot{\boldsymbol{p}}}{\partial \boldsymbol{p}} = -D\,N\,(\xi + \xi_P) \tag{5.116}$$
$$\frac{\partial \dot{\xi}}{\partial \xi} = 0 \tag{5.117}$$
$$\frac{\partial \dot{\eta}}{\partial \eta} = 0 \tag{5.118}$$
$$\frac{\partial \dot{V}}{\partial V} = D\,\xi_P \tag{5.119}$$
$$\frac{\partial \dot{\xi}_P}{\partial \xi_P} = 0 \tag{5.120}$$
$$\kappa(\boldsymbol{\eta}) \equiv -D\,\xi_P\,(N-1) - D\,N\,\xi \tag{5.121}$$

The metric determinant factor $\sqrt{g}$, introducing ϵ ($[\epsilon] = 1$), is:

$$\epsilon = \frac{1}{D}\,\ln\frac{V}{V_0} \tag{5.122}$$
$$\kappa(\boldsymbol{\eta}) = \dot{w}(\boldsymbol{\eta}) \tag{5.123}$$
$$\dot{w}(\boldsymbol{\eta}) = -D\,(N-1)\,\xi_P - D\,N\,\xi \tag{5.124}$$
$$w(\boldsymbol{\eta}) = -D\,(N-1)\,\epsilon - D\,N\,\eta \tag{5.125}$$
$$\sqrt{g(\boldsymbol{\eta})} = \mathrm{e}^{-w(\boldsymbol{\eta})} \tag{5.126}$$
$$\equiv \mathrm{e}^{D\,(N-1)\,\epsilon + D\,N\,\eta} \tag{5.127}$$

ϵ is inserted into equation (5.127), giving a volume prefactor.

$$\sqrt{g(\boldsymbol{\eta})} = \left(\frac{V}{V_0}\right)^{N-1} \mathrm{e}^{D\,N\,\eta} \tag{5.128}$$

Where V_0 is an arbitrary volume, usually the initial volume. Equation (5.122) shows that the barostat position ϵ is proportional to the logarithm of the ratio of two volumes. The time derivative of ϵ is ξ_P. ϵ is not a part of the equations of motion.

The partition function can now be calculated knowing the metric determinant factor and the conserved energy.

$$\begin{aligned}\Omega_{\text{NH NPT}} = \int \mathrm{d}^N\boldsymbol{p} \int_0^1 \mathrm{d}^N\boldsymbol{r} \int \mathrm{d}\xi \int \mathrm{d}\eta \int_0^\infty \mathrm{d}V \int \mathrm{d}\xi_P \left(\frac{V}{V_0}\right)^{N-1} \mathrm{e}^{D\,N\,\eta} \\ \times\delta\left[\mathcal{H}_0\left(\boldsymbol{p}, \boldsymbol{r}\,V^{\frac{1}{D}}\right) + \frac{1}{2}\,Q\,\xi^2 + X\,k_B\,T_E\,\eta + P_E\,V + \frac{1}{2}\,Q_P\,\xi_P^2 - E\right]\end{aligned} \tag{5.129}$$

An integration over η is done, leading to the final partition function generated by the Nosé-Hoover NPT equations of motion.

$$\eta_0 = \frac{E - \mathcal{H}_0\left(\boldsymbol{p}, \boldsymbol{r}\, V^{\frac{1}{D}}\right) - \frac{1}{2}\, Q\, \xi^2 - P_E\, V - \frac{1}{2}\, Q_P\, \xi_P^2}{X\, k_B\, T_E} \tag{5.130}$$

$$f'(\eta) = X\, k_B\, T_E \tag{5.131}$$

$$\begin{aligned} \Omega_{\text{NH NPT}} = & \frac{1}{X\, k_B\, T_E} \exp\left(\frac{D\, N}{X}\, \beta\, E\right) \\ & \times \int \mathrm{d}^N \boldsymbol{p} \int_0^1 \mathrm{d}^N \boldsymbol{r} \exp\left(-\frac{D\, N}{X}\, \beta\, \mathcal{H}_0\left(\boldsymbol{p}, \boldsymbol{r}\, V^{\frac{1}{D}}\right)\right) \\ & \times \int \mathrm{d}\xi \exp\left(-\frac{D\, N}{X}\, \frac{\beta}{2}\, Q\, \xi^2\right) \\ & \times \int_0^\infty \mathrm{d}V \left(\frac{V}{V_0}\right)^{N-1} \exp\left(-\frac{D\, N}{X}\, \beta\, P_E\, V\right) \\ & \times \int \mathrm{d}\xi_P \exp\left(-\frac{D\, N}{X}\, \frac{\beta}{2}\, Q_P\, \xi_P^2\right) \end{aligned} \tag{5.132}$$

Setting X equal to $D\, N$, the Nosé-Hoover partition function (5.133) is similar to the isothermal-isobaric partition function (5.134).

$$\Omega_{\text{NH NPT}} \propto \int_0^\infty \mathrm{d}V \left(\frac{V}{V_0}\right)^{N-1} \mathrm{e}^{-\beta\, P_E\, V} \int \mathrm{d}^N \boldsymbol{p} \int_0^1 \mathrm{d}^N \boldsymbol{r}\, \mathrm{e}^{-\beta\, \mathcal{H}_0\left(\boldsymbol{p}, \boldsymbol{r}\, V^{\frac{1}{D}}\right)} \tag{5.133}$$

$$\Omega_{\text{NPT}} = \frac{1}{N!\, h^{D\, N}\, V_0} \int_0^\infty \mathrm{d}V\, \mathrm{e}^{-\beta\, P_E\, V} \int \mathrm{d}^N \boldsymbol{p} \int_{D(V)} \mathrm{d}^N \boldsymbol{q}\, \mathrm{e}^{-\beta\, \mathcal{H}(\boldsymbol{p}, \boldsymbol{q})} \tag{5.134}$$

The volume term has an additional $\left(\frac{V}{V_0}\right)^{N-1}$ prefactor weighting the phase space. Whereas Nosé's NPT algorithm has a $\left(\frac{V}{V_0}\right)^{N}$ prefactor[37].

5.2.1. Choice of Q_P

The oscillation period, assuming harmonicity, is given by equation (5.135). Again, as seen in the thermostat mass section, a high mass decouples the barostat from the system, whereas a too low mass leads to higher fluctuations and inefficient phase space sampling. The rule of thumb is that the fluctuations should be the same order as a fundamental oscillation in the ensemble, *e.g.* an intramolecular vibration.

$$\tau(P) = \sqrt{\frac{Q_P}{D\, k_B\, T_E}} \tag{5.135}$$

The frequency does not depend on the system size.

5.2.2. Numerical Integration

For the numerical integration of the Nosé-Hoover equations of motion the modified Störmer version is used[34]. The algorithm is time reversible and highly stable.

With the standard integrators, Verlet or velocity Verlet, the differential equations cannot be solved in a direct way. Because of dependencies between the extended variables and the particles, the equations must be solved approximately by an iterative process. This is usually done with the Runge-Kutta or a predictor corrector algorithm. The disadvantage is that the integrator is no longer time reversible. The method discussed here uses the Verlet algorithm for the real particles (see equation (3.3) on page 22) and Störmer's centered-difference for the extended variables. The advantages are that the equations can be solved directly, the algorithm is time reversible, stable, accurate, fast and the memory usage is low.

The nomenclature introduced above in the Nosé-Hoover section will be used. A three dimensional space is assumed. Reduced coordinates are used which are similar to the internal variables in the simulation program used so far. The difference is that the reduced coordinates always have to be in the unit simulation cell (condition $0 \leq r_i < 1$) whereas the internal coordinates just have the dimension of 1 (position divided by the box length). The velocities are in unscaled SI units, $\mathrm{m\,s^{-1}}$.

The simulation program converts the coordinates into internal coordinates r' at the beginning:

$$\boldsymbol{r}' \;=\; \frac{\boldsymbol{q}}{BL(0)} \tag{5.136}$$

To use the NPT algorithm, reduced coordinates have to be calculated. This is done by moving the internal coordinates from the general space to the unit cell (equation (5.137)). The cell shifting distance can be easily calculated by the floor function, which is defined in the IEEE 754 floating point standard. It is available in Fortran 90 and C. The floor function returns the largest integer value which is not larger than the argument.

$$r \;=\; r' - \mathrm{floor}(r') \tag{5.137}$$

Value	Floor()
−99.1	−100
−99	−99
−0.1	−1
0.1	0
99	99

Table 5.1.: Floor Function Values.

Afterwards, the reduced coordinates are moved back to the original cell by adding the before subtracted shifting distance. This way vectors between coordinates of the same molecule can always be calculated by a simple difference. The atoms are close in the coordinate space, no cuts and inconsistencies are present. Distances between atoms are calculated by the minimum image convention. To calculate values in SI units, the distances have to be multiplied by the instantaneous box length. The space of the internal coordinates is an unscaled space and the scaling is done via the box length.

The Verlet algorithm for a general system is defined by equation (3.3):

$$\frac{\boldsymbol{r}_i(t+\delta t) - 2\,\boldsymbol{r}_i(t) + \boldsymbol{r}_i(t-\delta t)}{\delta t^2} = \ddot{\boldsymbol{r}}_i(t) \tag{5.138}$$

$\ddot{\boldsymbol{r}}_i(t)$ is calculated by the time derivative of equation (5.83).

$$\dot{\boldsymbol{r}}_i(t) = \frac{\boldsymbol{v}_i(t)}{V^{\frac{1}{3}}(t)} \tag{5.139}$$

$$\ddot{\boldsymbol{r}}_i(t) = \frac{\dot{\boldsymbol{v}}_i(t)\,V^{\frac{1}{3}}(t) - \boldsymbol{v}_i(t)\,\frac{1}{3}\,V^{-\frac{2}{3}}(t)\,\dot{V}(t)}{\left(V^{\frac{1}{3}}(t)\right)^2} \tag{5.140}$$

$$= \frac{\dot{\boldsymbol{v}}_i(t)}{V^{\frac{1}{3}}(t)} - \frac{1}{3}\,\boldsymbol{v}_i(t)\,V^{-\frac{4}{3}}(t)\,\dot{V}(t) \tag{5.141}$$

The time derivatives of the velocity and volume are given by the equations (5.84) and (5.86).

$$\dot{\boldsymbol{v}}_i(t) = \frac{\boldsymbol{f}_i(t)}{m_i} - (\xi + \xi_P)\,\boldsymbol{v}_i(t) \tag{5.142}$$

$$\dot{V}(t) = 3\,V(t)\,\xi_P(t) \tag{5.143}$$

Insertion into (5.141) gives:

$$\ddot{\boldsymbol{r}}_i(t) = \frac{\boldsymbol{f}_i(t)}{m_i\, V^{\frac{1}{3}}(t)} - \frac{(\xi(t) + \xi_P(t))\, \boldsymbol{v}_i(t)}{V^{\frac{1}{3}}(t)} - \frac{\boldsymbol{v}_i(t)\, \xi_P(t)}{V^{\frac{1}{3}}(t)} \tag{5.144}$$

$$= \frac{\boldsymbol{f}_i(t)}{m_i\, V^{\frac{1}{3}}(t)} - \frac{(\xi(t) + 2\,\xi_P(t))\, \boldsymbol{v}_i(t)}{V^{\frac{1}{3}}(t)} \tag{5.145}$$

Finally, the integration algorithm for the real particles is obtained by replacing the velocity with equation (3.4) and solving equation (5.146).

$$\boldsymbol{r}_i(t+\delta t) = 2\,\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-\delta t) + \left(\frac{\boldsymbol{f}_i(t)}{m_i\, V^{\frac{1}{3}}(t)} - \frac{(\xi(t) + 2\,\xi_P(t))\,(\boldsymbol{r}_i(t+\delta t) - \boldsymbol{r}_i(t-\delta t))}{2\,\delta t}\right) \delta t^2 \tag{5.146}$$

$$\boldsymbol{r}_i(t+\delta t) = \frac{2\,\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-\delta t)\left(1 - (\xi(t) + 2\,\xi_P(t))\,\frac{\delta t}{2}\right) + \frac{\boldsymbol{f}_i(t)}{m_i\, V^{\frac{1}{3}}(t)}\,\delta t^2}{1 + (\xi(t) + 2\,\xi_P(t))\,\frac{\delta t}{2}} \tag{5.147}$$

The extended variables are integrated by Störmer's central-difference:

$$\frac{V(t+\delta t) - V(t-\delta t)}{2\,\delta t} = \dot{V}(t) \tag{5.148}$$

$$= 3\,V(t)\,\xi_P(t) \tag{5.149}$$

$$\frac{\xi_P(t+\delta t) - \xi_P(t-\delta t)}{2\,\delta t} = \dot{\xi}_P(t) \tag{5.150}$$

$$= 3\,V(t)\,\frac{P(t) - P_E}{Q_P} \tag{5.151}$$

$$\frac{\xi(t+\delta t) - \xi(t-\delta t)}{2\,\delta t} = \dot{\xi}(t) \tag{5.152}$$

$$= \frac{2\,\mathcal{K}(t) - g\,k_B\,T_E}{Q} \tag{5.153}$$

P_E and T_E are the desired external pressure and temperature. $\mathcal{K}(t)$ is the kinetic energy of the real particles (equation (5.154)). $P(t)$ is the pressure given by the virial equation and the long-range correction (equation (5.155)). The long-range correction (equation (5.157)) is a function of the cut-off radius. If the cut-off radius is a function of the box length, then its value has to be recalculated each time step. The cut-off radius has to be smaller than half the box length.

$$\mathcal{K}(t) = \sum_{i=1}^{N} \frac{1}{2}\, m_i\, \boldsymbol{v}_i^2 \tag{5.154}$$

$$P(t) = \rho(t)\, k_B\, T + \frac{\mathcal{W}(t)}{V(t)} + P_{\mathrm{LR}}(t) \tag{5.155}$$

$$\mathcal{W} = \frac{1}{3} \sum_{i}^{N} \sum_{j>i}^{N} \boldsymbol{r}_{ij} \cdot \boldsymbol{f}_{ij} \tag{5.156}$$

$$P_{\mathrm{LR}} = -\frac{2}{3}\, \pi\, \rho^2 \int_{r_c}^{\infty} r^2\, w(r)\, \mathrm{d}\, r \tag{5.157}$$

$$w(r) = r\, \frac{\partial\, \Phi(r)}{\partial\, r} \tag{5.158}$$

All definitions stated so far are for atomic systems. If molecules are used, some additional points have to be considered. Atoms have three translational degrees of freedom whereas molecular systems have rotations and vibrations in addition. Often constraints are used to fix bond lengths and angles. The total degrees of freedom of a constrained system will be smaller than in the complete flexible case. The global thermostat is used in both cases with different values of g. For the pressure control, there exist two definitions of the pressure: the atomic pressure and the molecular pressure. Both have the same average value for an infinite system, but different fluctuations. The atomic pressure has fluctuations at higher frequencies due to the change of the intramolecular virial by intramolecular oscillations.

The atomic and molecular pressure are defined as:

$$P_{\mathrm{atm}} = \frac{\frac{2}{3}\, \mathcal{K} + \mathcal{W}}{V} + P_{\mathrm{LR}} \tag{5.159}$$

$$P_{\mathrm{mol}} = \frac{\frac{2}{3}\, \mathcal{K}_{\mathrm{mol}} + \mathcal{W}_{\mathrm{mol}}}{V} + P_{\mathrm{LR}} \tag{5.160}$$

The molecular pressure is defined by the kinetic energy of the molecule, the center of mass velocity plus the molecular virial which is defined by the forces on the center of mass between molecules.

$$\mathcal{K}_{\mathrm{mol}} = \sum_{i=1}^{N} \frac{1}{2}\, M_i\, \boldsymbol{v}_{i,\mathrm{mol}}^2 \tag{5.161}$$

$$\mathcal{W}_{\mathrm{mol}} = \frac{1}{3} \sum_{i}^{N} \sum_{j>i}^{N} \boldsymbol{R}_{ij} \cdot \boldsymbol{F}_{ij} \tag{5.162}$$

If constraints are used, then the constraint forces and the scaling of the coordinates are not independent. Two solutions exist for this problem[68]. The first is complicated and only solved for special cases by a few authors. The constraint forces are included into the integration algorithm[42, 15]. For simulations of a few and/or large molecules this method has to be used.

The disadvantage is that the algorithm is specific for one system and cannot be easily adapted to other problems. Furthermore, it was shown that several implementations do not sample the correct phase space or have hidden conditions[15].

The second solution separates the center of mass and intramolecular movement (*e.g.* in the `DL_PROTEIN` implementation[41]). The molecular pressure is used, which is independent of the intramolecular virial, the sum of all intramolecular forces is zero. Another advantage is that the same algorithm can be used for flexible molecules too. As described above, the molecular pressure has less fluctuations. Therefore the fluctuations of the volume due to the barostat are also smaller. Last but not least the algorithm is still time reversible whereas SHAKE is an iterative algorithm.

To simplify the equations, four additional values are defined which combine the thermostat and barostat variables of equation (5.147). R_f is the r-factor and D_f the division factor, the primed values are for the intramolecular movement without the barostat term.

$$R_f = 1 - \xi(t)\,\frac{\delta t}{2} - 2\,\xi_P(t)\,\frac{\delta t}{2} \tag{5.163}$$

$$D_f = 1 + \xi(t)\,\frac{\delta t}{2} + 2\,\xi_P(t)\,\frac{\delta t}{2} \tag{5.164}$$

$$R'_f = 1 - \xi(t)\,\frac{\delta t}{2} \tag{5.165}$$

$$D'_f = 1 + \xi(t)\,\frac{\delta t}{2} \tag{5.166}$$

The molecular force $\boldsymbol{F}_i$ acting on the center of mass is the sum of all atomic forces of the molecule. The new center of mass position of each molecule is given by:

$$\boldsymbol{R}_j(t+\delta t) = \frac{2\,\boldsymbol{R}_j(t) - \boldsymbol{R}_j(t-\delta t)\,R_f + \frac{\boldsymbol{F}_j(t)}{M_j\,V^{\frac{1}{3}}(t)}\delta t^2}{D_f} \tag{5.167}$$

The new atomic positions are calculated relative to the center of mass using a Nosé-Hoover thermostat:

$$\boldsymbol{r}_{ji}(t) = \boldsymbol{r}_i(t) - \boldsymbol{R}_j(t) \tag{5.168}$$

$$\boldsymbol{r}_{ji}(t-\delta t) = \boldsymbol{r}_i(t-\delta t) - \boldsymbol{R}_j(t-\delta t) \tag{5.169}$$

$$\boldsymbol{r}_i(t+\delta t) = \boldsymbol{R}_j(t+\delta t) + \frac{2\,\boldsymbol{r}_{ji}(t) - \boldsymbol{r}_{ji}(t-\delta t)\,R'_f + \left(\boldsymbol{f}_i(t) - \boldsymbol{F}_j(t)\,\frac{m_i}{M_j}\right)\frac{\delta t^2}{m_i\,V^{\frac{1}{3}}(t)}}{D'_f} \tag{5.170}$$

After that the velocities can be calculated via Verlet's equation (equation (3.4) on page 22).

The position of the thermostat η, which is necessary to calculate the total ensemble energy, but is not part of the integration algorithm, is integrated by the following equation:

$$\eta(t + \delta t) = \eta(t) + \xi(t)\,\delta t \tag{5.171}$$

The initial values of the extended variables at the time $t - \delta t$ have to be set to suitable values to guarantee a smooth start to the simulation. The first values at time t are:

$$\begin{aligned} \xi(t) &= 0 && (5.172)\\ \xi_P(t) &= 0 && (5.173)\\ \eta(t) &= 0 && (5.174) \end{aligned}$$

The initial volume is given by the density and the number of particles.

$$\rho(t) = \frac{N}{V(t)} \tag{5.175}$$

The values of the previous time step are calculated by the equations (5.148)-(5.153). A linear behavior is assumed:

$$\frac{V(t + \delta t) - V(t - \delta t)}{2\,\delta t} \approx \frac{V(t) - V(t - \delta t)}{\delta t} \tag{5.176}$$

Obtaining:

$$\begin{aligned} V(t - \delta t) &\approx -3\,V(t)\,\xi_P(t)\,\delta t + V(t) && (5.177)\\ &= V(t) && (5.178)\\ \xi(t - \delta t) &\approx -\frac{2\,\mathcal{K}(t) - g\,k_B\,T_E}{Q}\,\delta t + \xi(t) && (5.179)\\ &= -\frac{2\,\mathcal{K}(t) - g\,k_B\,T_E}{Q}\,\delta t && (5.180)\\ \xi_P(t - \delta t) &\approx -3\,V(t)\,\frac{P(t) - P_E}{Q_P}\,\delta t + \xi_P(t) && (5.181) \end{aligned}$$

$$= \quad -3\,V(t)\,\frac{P(t) - P_E}{Q_P}\,\delta t \tag{5.182}$$

The step by step implementation in the simulation program is:

1. Calculate the forces at time t: $\boldsymbol{f}(t)$
2. Calculate the next positions: $\boldsymbol{r}(t + \delta t)$
3. Calculate the velocity at time t: $\boldsymbol{v}(t)$
4. Calculate the long-range corrections at time t: $P_{\mathrm{LR}}(t)$
5. Maintain constraints (SHAKE): changes $\boldsymbol{r}(t + \delta t)$ and $\boldsymbol{v}(t)$
6. Calculate kinetic energy at time t: $\mathcal{K}(t)$
7. Calculate the pressure at time t: $P(t)$
8. Calculate the next volume: $V(t + \delta t)$
9. Calculate the next barostat rate: $\xi_P(t + \delta t)$
10. Calculate the next box length: $BL(t + \delta t)$
11. Scale intramolecular distances. Due to the changed box length, the intramolecular distances in reduced units have to be scaled by $BL(t + \delta t)/BL(t)$ to maintain the bond constraints.

    The same has to be done if the molecular pressure is used by the barostat. The intramolecular trajectory of a flexible molecule is independent of the barostat's volume scaling.
12. Calculate next thermostat rate: $\xi(t + \delta t)$
13. Calculate next thermostat position: $\eta(t + \delta t)$
14. Calculate next density: $\rho(t + \delta t)$
15. Calculate the conserved ensemble energy.
16. Increment the time step, rotate the variables and start from above.

6. Phase Equilibria

Different computational systems were developed to study the phase behavior of fluids[63]. In principle, several simulations at different phase points can be used to obtain the complete phase diagram. The fundamental problem is that in two-phase regions, both phases exist in the same system. Therefore a non-trivial interface region exists which occupies a large part of the simulation cell. Even with large systems, the phase separation is not stable enough to sample both single phases. Fluctuations of different magnitude and frequencies slow down the exchange of particles between the phases. However, the direct interface method is useful to study the structure of the interface.

6.1. Gibbs Ensemble Monte Carlo

The Gibbs Ensemble Monte Carlo method was introduced by Panagiotopoulos[61] in the year 1987. A μPT ensemble cannot be constructed because the variables are not independent. Therefore both phases have to be spatially separated into two regions. Periodic boundary conditions are used in both regions. The equilibrium condition is that both systems are in internal equilibrium and have an equal temperature, pressure and chemical potential for all components. Three different Monte Carlo moves are performed, in random order, to satisfy these conditions.

displacements	The particles are moved in each region according to the given external temperature (figure 6.1).
volume changes	The volume of each system fluctuates to reach the same pressure (figure 6.2). The total volume of all regions is constant.
particle transfers	Particles are transferred between the two systems to equalize the chemical potential (figure 6.3), using a Widom-like[95] process without knowing the exact value of the chemical potential.

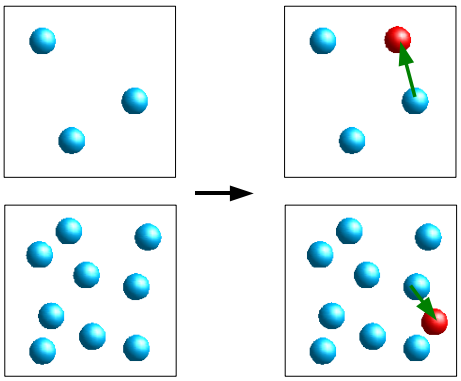


Figure 6.1.: Gibbs Ensemble Monte Carlo Displacement.

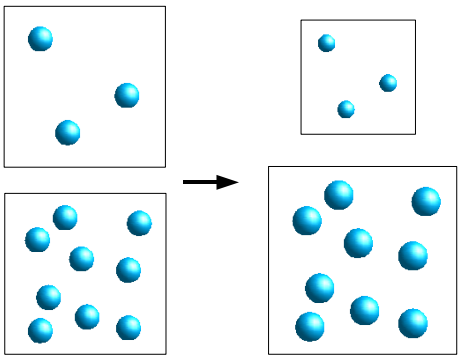


Figure 6.2.: Gibbs Ensemble Monte Carlo Volume Change.

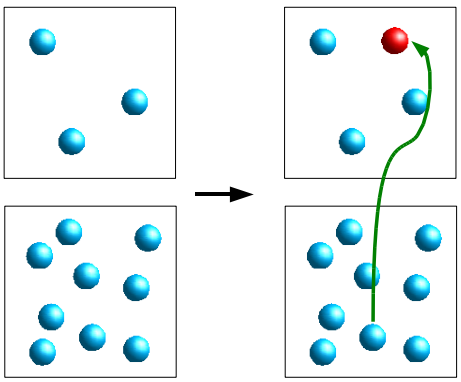


Figure 6.3.: Gibbs Ensemble Monte Carlo Particle Transfer.

The original NVT system is divided into two regions defined by the following quantities:

Property	**Region I**	**Region II**
Temperature:	T	T
Volume:	V_I	$V_{II} = V - V_I$
Particles:	N_I	$N_{II} = N - N_I$

The partition function Ω_{NVT} and the associated probability density $f(N_I, V_I; N, V, T)$ are

$$\begin{aligned} \Omega_{\mathrm{NVT}} &= \frac{1}{\Lambda^{3\,N}\,N!} \sum_{N_I=0}^{N} \binom{N}{N_I} \int_0^V \mathrm{d}V_I\, V_I^{N_I}\, V_{II}^{N_{II}} \int \mathrm{d}^{N_I}\boldsymbol{r}_I\, \mathrm{e}^{-\beta\, U_I(N_I)} \\ &\quad \times \int \mathrm{d}^{N_{II}}\boldsymbol{r}_{II}\, \mathrm{e}^{-\beta\, U_{II}(N_{II})} \qquad (6.1) \\ f(N_I, V_I; N, V, T) &\propto \frac{N!}{N_I!\, N_{II}!}\, \mathrm{e}^{N_I\, \ln V_I + N_{II}\, \ln V_{II} - \beta\, U_I(N_I) - \beta\, U_{II}(N_{II})} \qquad (6.2) \end{aligned}$$

where Λ is the thermal de Broglie wavelength.

The acceptance criteria for the three Monte Carlo moves are:

$$\begin{aligned} \wp_{\mathrm{move}} &= \min\left[1, \mathrm{e}^{-\beta\,\Delta U}\right] \qquad (6.3) \\ \wp_{\mathrm{volume}} &= \min\left[1, \mathrm{e}^{-\beta\,\Delta U_I - \beta\,\Delta U_{II} + N_I\, \ln\left(\frac{V_I + \Delta V}{V_I}\right) + N_{II}\, \ln\left(\frac{V_{II} - \Delta V}{V_{II}}\right)}\right] \qquad (6.4) \\ \wp_{\mathrm{transfer}} &= \min\left[1, \frac{N_{II}\, V_I}{(N_I + 1)\, V_{II}}\, \mathrm{e}^{-\beta\,\Delta U_I - \beta\,\Delta U_{II}}\right] \qquad (6.5) \end{aligned}$$

The displacements in equation (6.3) are done in each region. The volume V_I is increased by ΔV. The transfer in equation (6.5) takes place from region II to region I. The transfer probability from region I to region II is obtained by a circular rotation of the I and II indices. At first the source region of the transfer is determined and afterwards the transfer particle. The Metropolis algorithm is used to decide if a move is acceptable (see Monte Carlo chapter on page 26).

Gibbs' phase rule[7]

$$F = c - p + 2 \qquad (6.6)$$

(where c is the number of components, p the number of phases and F the degrees of freedom) gives the smallest number of intensive variables which must be defined to completely describe

a state of the system. For a single component system ($c = 1$) with two phases ($p = 2$), there exists one degree of freedom. In the Gibbs ensemble method, the temperature is the degree of freedom which defines the complete system.

The Gibbs ensemble method can be used for pure fluids and mixtures. The algorithm works up to temperatures close to the critical temperature, but in the near critical region long-range effects play a role and even large systems cannot sample these states. Another problem is that the free-energy cost to have both phases in one region gets smaller. Away from the critical temperature, the free-energy gain is the driving factor of the phase separation.

Gibbs ensemble Monte Carlo simulations are started at a certain temperature and two densities in the two-phase region. The simulations are performed until all three equilibrium conditions are fulfilled: equal temperature, equal pressure and stable phase densities.

6.1.1. Limits

The particle transfer probability strongly depends on the density in both regions. Moving to higher densities (lower temperatures) decreases the exchange probability and the time to reach equilibrium is increased.

Another problem inherent to the Monte Carlo method is that molecular systems are much more complicated to sample than simple atomic systems. Random positions for additional degrees of freedom have to be generated. The orientation of the molecules and possible intramolecular movements are sampled by additional Monte Carlo steps. Molecular systems also increase the complexity of the particle transfer step. The probability to find a "hole" in a dense phase is several times smaller than for an atom which is represented by a point, but a certain percentage of the transfers have to be successful to reach the equilibrium condition. Several extensions were developed to improve the acceptance for molecules. One method for multi-segment molecules, *e.g.* *n*-alkanes and polymers, the "configurational-bias sampling" transfers one segment after the other[24]. The first position is at random coordinates, the next segment position is selected by trying several growth directions. At the end an overall probability is calculated which is used as the acceptance factor in the Metropolis algorithm. A special case of the "configurational-bias sampling" is the "rotational insertion bias method"[17]. An insertion point is calculated and afterwards several orientations of the molecule are tried to find an optimal geometry. This method is most efficient for small molecules. Gibbs ensemble simulations which use this method were done for systems of water[18] and CO_2. The water results depend strongly on the potential and the critical temperatures vary by 50 Kelvin. The CO_2 EPM2[31, 93] potential and most other potentials were fitted to thermodynamic properties to reproduce the experimental critical point. Therefore these results are more a check on how well the potential scales in the liquid-vapor region than a real calculation of the critical properties.

No trajectories or time dependent values can be obtained from the Monte Carlo steps since only a random walk through the configuration space is calculated. The momenta of the particles are unknown and therefore the pressure has to be calculated by using other properties, for example the energy fluctuations due to the volume change.

6.2. Gibbs Ensemble Molecular Dynamics

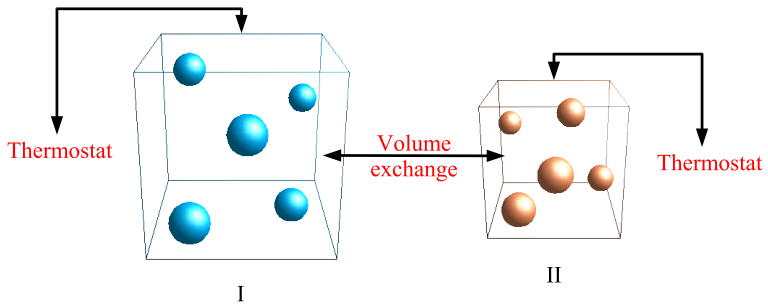
The limiting factor of the Gibbs ensemble Monte Carlo method is the low probability of the particle transfer. This problem can be solved by using a molecular dynamics approach instead of the Monte Carlo sampling.

The newly developed algorithm for the Gibbs ensemble molecular dynamics method is similar to Palmer and Lo's approach[60], the first molecular dynamics algorithm for the Gibbs ensemble. The temperature and pressure are controlled by enhanced equations of motion. The chemical potential is equalized by a hybrid Monte Carlo-molecular dynamics algorithm. The main differences to Palmer and Lo are that Hoover's non-Hamiltonian dynamics is used (equations (6.56)-(6.69)), which avoids the scaling of the time, the exchange Hamiltonian $\mathcal{H}_{\text{xch}}$ (equation (6.17)) is modified for a smoother particle exchange and the algorithm was extended from atomic Lennard-Jones potentials to molecular potentials of any shape (see figure 6.4). The state of the art non-Hamiltonian theory (see section 2.3 on page 16) is used to prove that the correct partition function is sampled and that the particle exchange equalizes the chemical potential.

In the following sections L is the number of coupled subsystems (usually two). A subsystem in this context is a system of N_i molecules in a cubic box of volume V_i. Periodic boundary conditions are applied to all subsystems. Roman numbers are used to label the subsystems.

Two liquid-gas systems starting in the two-phase region will separate into two phases. The algorithm presented here works unmodified for liquid-gas phases. The study of solid phases often needs additional constraints and parameters to guide the insertion to unoccupied places in the crystal structure. An example of a solid-solid phase transition is given in reference [23].

6.2.1. Equal Temperature



The thermalization is done separately for each individual phase using L Nosé-Hoover thermostats (see section 5.1.2 on page 32). The thermostat frequency (equation (5.79) on page 38) depends on the degrees of freedom, therefore the thermostat mass Q is scaled by the degrees of freedom to obtain a constant frequency in all phases. At the initialization of the program, the highest frequency of all subsystems is used to calculate τ_0.

$$\tau_0 = 2\,\pi\sqrt{\frac{Q_0}{g_0\,k_B\,T_E}} \tag{6.7}$$

$$Q_i(t) = \frac{\tau_0^2}{4\,\pi^2}\,g(t)\,k_B\,T_E \tag{6.8}$$

$$= \frac{Q_0}{g_0}\,g(t) \tag{6.9}$$

The constant ratio Q_0/g_0 is the "thermostat mass per degree of freedom".

6.2.2. Equal Pressure

The Nosé-Hoover isothermal-isobaric ensemble couples one system to an external pressure. To equalize the pressure between different phases, the barostat is coupled to the average pressure of all phases.

$$\bar{P}(t) = \frac{1}{L}\sum_{i=1}^{L} P_i(t) \tag{6.10}$$

$$V = \sum_{i=1}^{L} V_i(t) \tag{6.11}$$

$$= \text{const} \tag{6.12}$$

In a two-phase system, the barostat is coupled to the first subsystem. The volume change of the first subsystem is given by the Nosé-Hoover equations, the volume change of the second subsystem follows by the conservation of the total volume. Exchanging the subsystems would lead to the same dynamics.

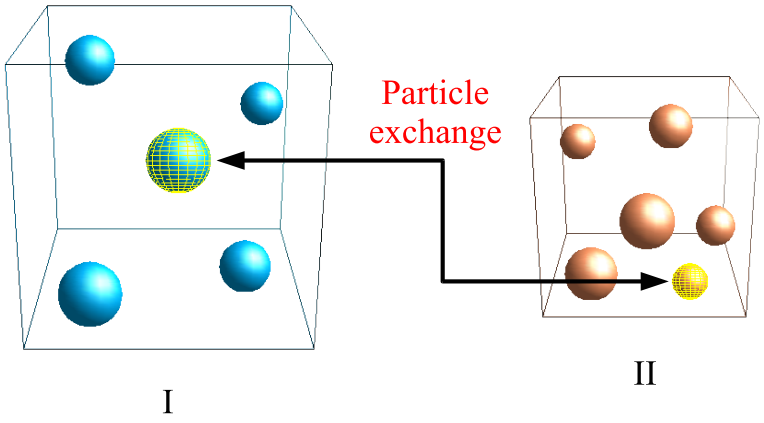
For more than two phases, the following algorithm could be used. $L-1$ subsystems are coupled to a barostat, the last subsystem acts as a volume reservoir. This algorithm is implemented in the simulation program but was never used.

Periodic Boundary Conditions

Periodic boundary conditions are used in both subsystems, decreasing the degrees of freedom by D in all subsystems. The total linear momentum has to be zero in all subsystems. Without particle exchanges, the linear momentum is conserved. With exchanges, the momenta flow introduced by the fractional particle has to be corrected. See the particle exchange section 6.2.3 for more details.

A cut-off radius is used in each subsystem. For equal conditions, the cut-off radius has to be the same in all subsystems. As a consequence, the long-range correction of the potential and virial have no influence on the equilibrium, they have the same value everywhere at the same time. Therefore only the short-range interactions determine the equilibrium conditions. It is important to use an appropriate cut-off value.

6.2.3. Equal Chemical Potential



The Gibbs ensemble Monte Carlo method is limited to a single step transfer. In contrast, the molecular dynamics equations of motion generate a trajectory in the phase space. Therefore it is possible to introduce a multi time step algorithm for the particle exchange, having the advantage that the exchange probability can be optimized and complicated molecular systems or dense phases are no longer a limit. The exchange is a combination of an initial Monte Carlo step to define the first exchange parameters and a one dimensional exchange equation of motion. Due to the Monte Carlo step, the conserved energy given by the molecular dynamics pseudo Hamiltonian (6.95) is no longer constant, the particle exchange is an unphysical process. In steps without exchanges and during transfers, the Hamiltonian is conserved and the algorithm guarantees that the temperatures and pressures are constant. The transfer particle is called the fractional particle. The degree of transfer between the source and destination subsystems is given by the position ν ($[\nu] = 1$), where ν is in the range [0,1]. A value of one means that the fractional particle is fully integrated into the subsystem and is equal to a normal particle, a value of zero means that the fractional particle is not part of the subsystem and has no interactions with normal particles. There exists only one fractional particle at any given time step. To describe interactions between normal particles and the fractional particle, a modified potential function Φ_f is used. The algorithm shown here is one possibility to exchange particles to reach equal chemical potentials. For other systems, the algorithm could be enhanced or replaced.

The shifted potential used between the fractional particle and real particles is:

$$\Phi_f = \nu \sum_{i=1}^{N} \Big(\Phi(q'_{if}) - \Phi(q_{cf}) \Big) \tag{6.13}$$

The fractional potential is shifted to have a value of zero at the cut-off radius r_c. The distances between the fractional molecule and the real molecules are increased to avoid singularities and high repulsive forces at short distances. This means that the volume occupied by the fractional particle in each subsystem is related to its ν.

$$q'_{if} = q_{if} + (1-\nu)\,\eta \tag{6.14}$$

$$\boldsymbol{q}'_{if} = \boldsymbol{q}_{if} + \frac{\boldsymbol{q}_{if}}{q_{if}}\,(1-\nu)\,\eta \tag{6.15}$$

$$q_{cf} = r_c + (1-\nu)\,\eta \tag{6.16}$$

η ($[\eta] = \mathrm{m}$) is the maximal increase in distance. For a further discussion of η see the parameter section on page 65.

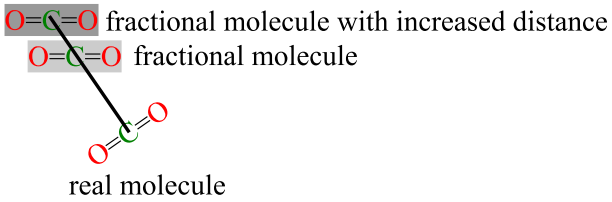


Figure 6.4.: Fractional Particle Distance Increase.

During the calculation of the forces and the higher derivatives including the virial contributions, the distance and the distance vector between two centers of mass of the monomers is evaluated and modified according to equation (6.14) and (6.15). A problem would arise if both centers are the same, leading to an undefined vector in equation 6.15. The probability of this situation is very low and did not happen in several hundred million simulation steps. Anyway, a further security check was added to the program. If a new molecule is inserted in a simulation box, the minimal allowed center of mass distance is guaranteed to be larger than 10 pm.

The equations of motion of the exchange coordinate ν are based on a Hamiltonian (6.17) which was proposed by Palmer and Lo. In contrast to their work, the mass is also scaled by the exchange position ν to allow a smooth transfer of the molecule, including its momenta (see equations (6.22) and (6.23) and the first terms of (6.18) and (6.19)).

$$\mathcal{H}_{\mathrm{xch}} = \mathcal{H}_I + \mathcal{H}_{II} + \mathcal{H}_\nu \tag{6.17}$$

$$\mathcal{H}_I = \frac{\boldsymbol{p}'^2_{f,I}}{2\,m_f\,\nu_I} + \Phi_{f,I}(\boldsymbol{q}_I, \boldsymbol{q}_{f,I}, \nu_I) \tag{6.18}$$

$$\mathcal{H}_{II} = \frac{\boldsymbol{p}'^2_{f,II}}{2\,m_f\,\nu_{II}} + \Phi_{f,II}(\boldsymbol{q}_{II}, \boldsymbol{q}_{f,II}, \nu_{II}) \tag{6.19}$$

$$\begin{aligned}\mathcal{H}_\nu &= \frac{p^2_{\nu_I}}{2\,Q_E} + \nu_I\,k_B\,T_E\,\ln((N_I+1)!\,N_{II}!\,V_{II}) \\ &\quad +\nu_{II}\,k_B\,T_E\,\ln((N_{II}+1)!\,N_I!\,V_I) \qquad (6.20)\\ &= \frac{p^2_{\nu_I}}{2\,Q_E} + \nu_I\,k_B\,T_E\,\ln\frac{(N_I+1)\,V_{II}}{(N_{II}+1)\,V_I} + k_B\,T_E\,\ln((N_{II}+1)!\,N_I!\,V_I) \qquad (6.21)\end{aligned}$$

$$\boldsymbol{p}'_{f,I} = m_f\,\nu_I\,\dot{\boldsymbol{q}}_{f,I} \tag{6.22}$$

$$\boldsymbol{p}'_{f,II} = m_f\,\nu_{II}\,\dot{\boldsymbol{q}}_{f,II} \tag{6.23}$$

$$\nu_I + \nu_{II} = 1 \tag{6.24}$$

Where m_f is the mass of the fractional particle, ν the fractional position, p_ν its conjugated momentum and Q_E ($[Q_E] = \mathrm{kg\,m^2}$) the exchange mass.

The equations of motion are obtained in the usual way.

$$\dot{\nu}_I = \frac{\partial\,\mathcal{H}_{\mathrm{xch}}}{\partial\,p_{\nu_I}} \tag{6.25}$$

$$= \frac{p_{\nu_I}}{Q_E} \tag{6.26}$$

$$\dot{p}_{\nu_I} = -\frac{\partial\,\mathcal{H}_{\mathrm{xch}}}{\partial\,\nu_I} \tag{6.27}$$

$$\begin{aligned}&= -\frac{\boldsymbol{p}'^2_{f,I}}{2\,m_f\,\nu_I^2} - \frac{\partial\,\Phi_{f,I}(\boldsymbol{q}_I, \boldsymbol{q}_{f,I}, \nu_I)}{\partial\,\nu_I} + \frac{\boldsymbol{p}'^2_{f,II}}{2\,m_f\,\nu_{II}^2} + \frac{\partial\,\Phi_{f,II}(\boldsymbol{q}_{II}, \boldsymbol{q}_{f,II}, \nu_{II})}{\partial\,\nu_{II}} \\ &\quad +k_B\,T_E\,\ln\frac{(N_{II}+1)\,V_I}{(N_I+1)\,V_{II}} \qquad (6.28)\end{aligned}$$

The relation $\mathrm{d}\nu_I = -\mathrm{d}\nu_{II}$ and therefore $\frac{\partial\,\Phi_f(\nu_{II})}{\partial\,\nu_I} = -\frac{\partial\,\Phi_f(\nu_{II})}{\partial\,\nu_{II}}$ were used to calculate the expression for the "fractional force" $\dot{p}_\nu$. The "fractional force" has the same units as the potential. The forces acting on the fractional molecule in both subsystems are defined as

$$f_{\nu_I} = -\frac{\partial\,\Phi_{f,I}(\boldsymbol{q}_I, \boldsymbol{q}_{f,I}, \nu_I)}{\partial\,\nu_I} \tag{6.29}$$

$$f_{\nu_{II}} = -\frac{\partial\,\Phi_{f,II}(\boldsymbol{q}_{II}, \boldsymbol{q}_{f,II}, \nu_{II})}{\partial\,\nu_{II}} \tag{6.30}$$

which simplifies equation (6.28) to:

$$\dot{p}_{\nu_I} = -\frac{\boldsymbol{p}'^2_{f,I}}{2\,m_f\,\nu_I^2} + \frac{\boldsymbol{p}'^2_{f,II}}{2\,m_f\,\nu_{II}^2} + f_{\nu_I} - f_{\nu_{II}} + k_B\,T_E\,\ln\frac{(N_{II}+1)\,V_I}{(N_I+1)\,V_{II}} \tag{6.31}$$

Transforming the kinetic energy terms from the Hamiltonian space to the Lagrangian space leads to an unscaled equation (6.32). The terms are the unscaled kinetic energy of the exchange molecule. With a zero ν_I or ν_{II} the Lagrangian terms vanish in the Hamiltonian. On the other hand, the term in the Hamiltonian is a "division by zero". Palmer and Lo avoided this problem by leaving the mass scaling away and assumed that the integration over the momenta generates a constant. The latter is not true: the momenta are scaled by the thermostat and barostat. There exists another problem: the momentum introduced in the first exchange step has to be removed to obey the momentum conservation law in each subsystem. If the complete momentum is added on each insertion and removed on each failed exchange, a considerable unphysical error is introduced. In the algorithm which is presented here, only a fraction of the total momentum flows into the system. If the fractional particle exists for a few time steps then the momentum flow is negligible. The same is true for the instantaneous momentum correction, because the momentum flow between time steps is small, the unphysical correction is minimized.

$$\frac{\boldsymbol{p}'^2_f}{2\,m_f\,\nu^2} = \frac{1}{2}\,m_f\,\dot{\boldsymbol{q}}_f^2 \tag{6.32}$$

Analyzing the "division by zero" with l'Hospitals rule[52], shows that the second derivative is equal to the Lagrangian value. Therefore the Hamiltonian value is always equal to the Lagrangian one. The numerical problem in the computer code can completely be avoided by using the generalized coordinates instead of the generalized momenta.

$$\lim\frac{f(\nu)}{g(\nu)} = \lim\frac{f'(\nu)}{g'(\nu)} \tag{6.33}$$

$$= \lim\frac{f''(\nu)}{g''(\nu)} \tag{6.34}$$

$$f(\nu) = m_f^2\,\nu^2\,\dot{\boldsymbol{q}}_f^2 \tag{6.35}$$

$$g(\nu) = 2\,m_f\,\nu^2 \tag{6.36}$$

$$f''(\nu) = 2\,m_f^2\,\dot{\boldsymbol{q}}_f^2 \tag{6.37}$$

$$g''(\nu) = 4\,m_f \tag{6.38}$$

$$\lim\frac{f(\nu)}{g(\nu)} = \frac{1}{2}\,m_f\,\dot{\boldsymbol{q}}_f^2 \tag{6.39}$$

Separating the fractional force $\dot{p}_\nu$ into independent equations for each subsystem gives these equations.

$$\dot{p}_{\nu_I} = \dot{p}_{\nu_I}(I) - \dot{p}_{\nu_I}(II) \tag{6.40}$$

$$\dot{p}_{\nu_I}(I) = -\mathcal{K}_{f,I} + f_{\nu_I} - k_B\,T_E \ln \frac{N_I + 1}{V_I} \tag{6.41}$$

$$\dot{p}_{\nu_I}(II) = -\mathcal{K}_{f,II} + f_{\nu_{II}} - k_B\,T_E \ln \frac{N_{II} + 1}{V_{II}} \tag{6.42}$$

The number of molecules is no longer an integer value, it is defined by the sum of the real molecules plus the fractional position of the exchange particle. All other N dependent variables have to be adjusted too. The expression for the molar density is given in equation (6.44).

$$N_i'(t) = N_i(t) + \nu_i(t) \tag{6.43}$$

$$\rho_i'(t) = \frac{N_i'(t)}{V_i(t)\,N_A} \tag{6.44}$$

Linear Momentum

As pointed out before, the total momentum is a conserved quantity of the equations of motion. If periodic boundary conditions are used, then the total momentum has to be zero to generate the correct isothermal-isobaric distribution. Here these conditions have to be fulfilled in each subsystem. The equations given next are for subsystem I.

$$\sum_{i=1}^{N_I} \boldsymbol{p}_i + \boldsymbol{p}'_{f,I} \equiv 0 \tag{6.45}$$

Advancing the exchange from $\nu_I(0)$ to $\nu_I(1)$ adds the momentum $m_f\,\dot{\boldsymbol{q}}_{f,I}(\nu_I(1) - \nu_I(0))$. Therefore the velocities of the real particles are shifted to obtain a total momentum of zero. Here it will be shown how the correction is done for the Verlet integrator.

The previous positions $\boldsymbol{r}(t)$, $\boldsymbol{r}(t-\delta t)$ and velocities $\boldsymbol{v}(t-\delta t)$ fulfilled condition (6.45). The Verlet integrator gives new positions $\boldsymbol{r}(t+\delta t)$ and velocities $\boldsymbol{v}(t)$ with a non-zero total momentum $\boldsymbol{p}_{\mathrm{box}}(t)$.

$$\boldsymbol{p}_{\mathrm{box}}(t) = \sum_{i=1}^{N_I} m_i\,\boldsymbol{v}_i(t) + m_f\,\nu_I(t)\,\boldsymbol{v}_{f,I}(t) \tag{6.46}$$

The positions $\boldsymbol{r}(t-\delta t)$ are treated as constant, therefore the newer positions are shifted. The correction of the velocity is given by:

$$\boldsymbol{v}_{\mathrm{corr}}(t) = \frac{\boldsymbol{p}_{\mathrm{box}}(t)}{\sum_{i=1}^{N_I} m_i} \tag{6.47}$$

$$\boldsymbol{v}'(t) = \boldsymbol{v}(t) - \boldsymbol{v}_{\mathrm{corr}}(t) \tag{6.48}$$

If both following equations are inserted into the velocity correction equation (6.48)

$$\boldsymbol{v}(t) = \frac{\boldsymbol{r}(t+\delta t) - \boldsymbol{r}(t-\delta t)}{2\,\delta t}\, BL(t) \tag{6.49}$$

$$\boldsymbol{v}'(t) = \frac{\boldsymbol{r}'(t+\delta t) - \boldsymbol{r}(t-\delta t)}{2\,\delta t}\, BL(t) \tag{6.50}$$

one obtains the correction for the next positions:

$$\boldsymbol{r}'(t+\delta t) = \boldsymbol{r}(t+\delta t) - \boldsymbol{v}_{\mathrm{corr}}(t)\,\frac{2\,\delta t}{BL(t)} \tag{6.51}$$

It is important that these corrections are only done for the real particles, not for the fractional particle. The fractional particle is the source of the non-zero total momentum, but not affected by the momentum rescaling. The correction changes the positions between two time steps, and hence the velocity, but not the distances between particles of the same time step.

The positions at time t have to be similarly corrected. This is done by the calculation of the half step velocity and the momentum correction.

$$\boldsymbol{v}\left(t+\frac{\delta t}{2}\right) = \frac{\boldsymbol{r}(t+\delta t) - \boldsymbol{r}(t)}{\delta t}\, BL(t) \tag{6.52}$$

$$\boldsymbol{p}_{\mathrm{box}}\left(t+\frac{\delta t}{2}\right) = \sum_{i=1}^{N_I} m_i\, \boldsymbol{v}_i\left(t+\frac{\delta t}{2}\right) + m_f\, \nu_I(t)\, \boldsymbol{v}_{f,I}\left(t+\frac{\delta t}{2}\right) \tag{6.53}$$

$$\boldsymbol{v}_{\mathrm{corr}}\left(t+\frac{\delta t}{2}\right) = \frac{\boldsymbol{p}_{\mathrm{box}}\left(t+\frac{\delta t}{2}\right)}{\sum_{i=1}^{N_I} m_i} \tag{6.54}$$

$$\boldsymbol{r}'(t) = \boldsymbol{r}(t) - \boldsymbol{v}_{\mathrm{corr}}\left(t+\frac{\delta t}{2}\right)\,\frac{\delta t}{BL(t)} \tag{6.55}$$

Algorithm

The exchange algorithm is shown in the following graph. At each step three different states are possible: there is no exchange in this step, an exchange is already running or an exchange will be initialized. Between exchanges, a random number of steps, no new exchange will be started. This is done to give the system time to equilibrate and to control the exchange rate.

ν_0 is the initial position of the particle. In all simulations used in this work it has the value zero. On the other hand v_0 is the initial velocity of the exchange. v_0 has the dimension of 1 and is the increment (or decrement) of the position ν per exchange time step. If no fractional forces act on the exchange particle, then the exchange would be finished in $1/v_0$ time steps. v_0 is Gaussian distributed with a variance of 10 % to avoid systematic errors due to fixed exchange lengths.

Each of the exchange steps is done in one simulation step, no multi time-step algorithm is used.

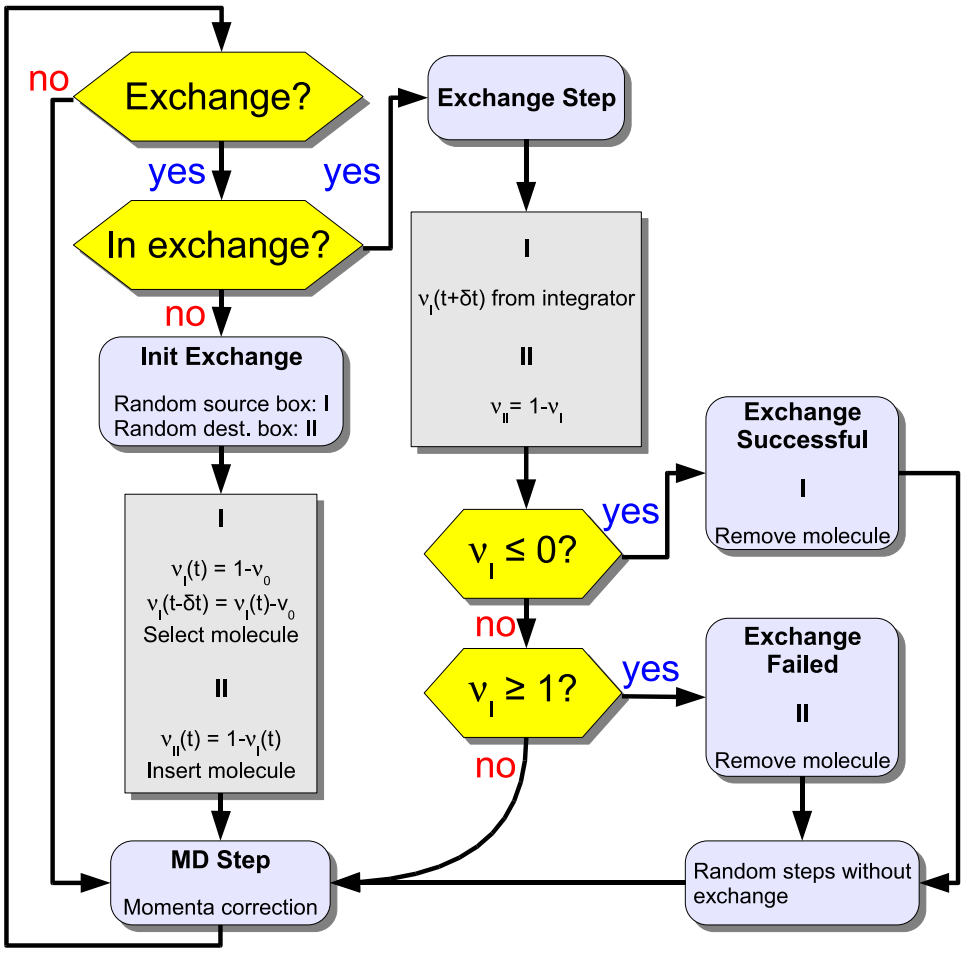


Figure 6.5.: GEMD Exchange Algorithm Scheme

To insert a new particle into a system, several initial properties have to be defined. First the center of mass position ($0 \leq r_{\mathrm{CM}} < 1$) and the orientation of the molecule (three Euler angles) have to be set. The random number generator is used to calculate these values. The molecular velocities and positions are separated into translations, rotations and intramolecular oscillations. A random phase is used to set the velocities and displacements of each normal mode (see appendix B.1.2 on page 133). For the translation and rotation components, values for 100 molecules are calculated in advance and normalized to a Gaussian/Boltzmann distribution. After 100 particle creations, all values have been used and a new set of 100 values is generated.

Parameters

The optimal values for the exchange mass Q_E, the initial velocity v_0 and the maximal distance increase η depend on the chosen system and potential. They have to be adjusted to reach equal chemical potentials in an efficient way. Some rule of thumb how to get these values will be shown next.

The maximal distance increase η of the potential must be larger than the diameter of the molecule, otherwise atoms of the fractional particle and normal particles can overlap, leading to singularities and high energies. The maximum η value should not be larger than the cut-off radius. Too high η values lead to a slower convergence of the chemical potential. The optimal value of η can be obtained from a graphical plot of different dimer orientations (see figure 7.1 on page 74). Above most of the highly repulsive distances are the optimal values.

The initial velocity v_0 is the inverse of the number of time steps an exchange would take without fractional forces. The exchange mass Q_E determines the strength of the fractional force coupling. A large Q_E value gives a system where particles are transferred between the subsystem in $1/v_0$ time steps. In this case, the exchanges have no influence on the chemical potential. If Q_E is too small, most particle transfers will fail in a few steps and successful transfers will be done too quickly. The transfer is inefficient and it is possible that the internal equilibrium in the subsystems is destroyed. Both parameters are not independent, the exchange length depends on the ability of the target subsystem to integrate the fractional particle. Rotational and translational reorganization times play an important role. The easiest way is to start with an appropriate value of v_0, equal to the reorganization time. Then the exchange mass Q_E is changed to find the point where the influence of the fractional potential begins. Below this Q_E value, the number of failed exchanges will increase and the length of failed exchanges gets smaller. To be in the optimal region, the failed length should not be too small, but not larger than successful exchanges.

6.2.4. Partition Function and Exchange Probability

The equations of motion for a two-phase system are:

$$\dot{\boldsymbol{r}}_{I,i} = \frac{\boldsymbol{p}_{I,i}}{m_{I,i}\, V_I^{\frac{1}{D}}} \tag{6.56}$$

$$\dot{\boldsymbol{r}}_{II,i} = \frac{\boldsymbol{p}_{II,i}}{m_{II,i}\, V_{II}^{\frac{1}{D}}} \tag{6.57}$$

$$\dot{\boldsymbol{r}}_{f,I} = \frac{\boldsymbol{p}_{f,I}}{m_f\, V_I^{\frac{1}{D}}} \tag{6.58}$$

$$\dot{\boldsymbol{r}}_{f,II} = \frac{\boldsymbol{p}_{f,II}}{m_f\, V_{II}^{\frac{1}{D}}} \tag{6.59}$$

$$\dot{\boldsymbol{p}}_{I,i} = -\frac{\partial\, \Phi(\boldsymbol{q}_I)}{\partial\, \boldsymbol{q}_{I,i}} - \frac{\partial\, \Phi_f(\boldsymbol{q}_I, \boldsymbol{q}_{f,I}, \nu_I)}{\partial\, \boldsymbol{q}_{I,i}} - (\xi_P + \xi_I)\, \boldsymbol{p}_{I,i} \tag{6.60}$$

$$\dot{\boldsymbol{p}}_{II,i} = -\frac{\partial\, \Phi(\boldsymbol{q}_{II})}{\partial\, \boldsymbol{q}_{II,i}} - \frac{\partial\, \Phi_f(\boldsymbol{q}_{II}, \boldsymbol{q}_{f,II}, \nu_{II})}{\partial\, \boldsymbol{q}_{II,i}} - \xi_{II}\, \boldsymbol{p}_{II,i} \tag{6.61}$$

$$\dot{\boldsymbol{p}}_{f,I} = -\frac{\partial\, \Phi_f(\boldsymbol{q}_I, \boldsymbol{q}_{f,I}, \nu_I)}{\partial\, \boldsymbol{q}_{f,I}} - (\xi_P + \xi_I)\, \boldsymbol{p}_{f,I} \tag{6.62}$$

$$\dot{\boldsymbol{p}}_{f,II} = -\frac{\partial\, \Phi_f(\boldsymbol{q}_{II}, \boldsymbol{q}_{f,II}, \nu_{II})}{\partial\, \boldsymbol{q}_{f,II}} - \xi_{II}\, \boldsymbol{p}_{f,II} \tag{6.63}$$

$$\dot{\xi}_I = \frac{\sum_{i=1}^{N_I} \frac{\boldsymbol{p}_{I,i}^2}{m_{I,i}} + \frac{\boldsymbol{p}_{f,I}'^2}{m_f\, \nu_I} - g_I\, k_B\, T_E}{Q_I} \tag{6.64}$$

$$\dot{\xi}_{II} = \frac{\sum_{i=1}^{N_{II}} \frac{\boldsymbol{p}_{II,i}^2}{m_{II,i}} + \frac{\boldsymbol{p}_{f,II}'^2}{m_f\, \nu_{II}} - g_{II}\, k_B\, T_E}{Q_{II}} \tag{6.65}$$

$$\dot{V}_I = D\, V_I\, \xi_P \tag{6.66}$$

$$\dot{\xi}_P = \frac{D\, V_I (P_I - \bar{P})}{Q_P} \tag{6.67}$$

$$\dot{\nu}_I = \frac{p_{\nu_I}}{Q_E} \tag{6.68}$$

$$\dot{p}_{\nu_I} = -\frac{\boldsymbol{p}_{f,I}'^2}{2\, m_f\, \nu_I^2} - \frac{\partial\, \Phi_f(\boldsymbol{q}_I, \boldsymbol{q}_{f,I}, \nu_I)}{\partial\, \nu_I} + \frac{\boldsymbol{p}_{f,II}'^2}{2\, m_f\, \nu_{II}^2} + \frac{\partial\, \Phi_f(\boldsymbol{q}_{II}, \boldsymbol{q}_{f,II}, \nu_{II})}{\partial\, \nu_{II}} + k_B\, T_E\, \ln \frac{(N_{II}+1)\, V_I}{(N_I+1)\, V_{II}} \tag{6.69}$$

The coordinates $\boldsymbol{r}_i$ are independent of the volume and are in the range $0 \leq r_i < 1$. The barostat velocity and force act on subsystem I and the volume of subsystem II is calculated by the conserved total volume. The time derivative of the second volume is:

$$\dot{V}_{II} = -\dot{V}_I \tag{6.70}$$

The fractional particle follows the same equations of motion in the subsystem as a real particle.

The force of the real molecules in the subsystem acts on the complete fractional molecule. The potential, momenta, virial and other interactions are scaled by ν.

The compressibility $\kappa(\boldsymbol{\eta})$ of the non-Hamiltonian phase space is:

$$\begin{aligned}
\kappa(\boldsymbol{\eta}) &= \boldsymbol{\nabla}_{\boldsymbol{\eta}} \cdot \dot{\boldsymbol{\eta}} && (6.71)\\
&= \sum_{i=1}^{2D(N_I+N_{II}+\nu_I+\nu_{II})+8} \boldsymbol{\nabla}_{\boldsymbol{\eta}_i} \cdot \dot{\boldsymbol{\eta}}_i && (6.72)\\
\frac{\partial \dot{\boldsymbol{r}}_I}{\partial \boldsymbol{r}_I} &= 0 && (6.73)\\
\frac{\partial \dot{\boldsymbol{r}}_{II}}{\partial \boldsymbol{r}_{II}} &= 0 && (6.74)\\
\frac{\partial \dot{\boldsymbol{r}}_{f,I}}{\partial \boldsymbol{r}_{f,I}} &= 0 && (6.75)\\
\frac{\partial \dot{\boldsymbol{r}}_{f,II}}{\partial \boldsymbol{r}_{f,II}} &= 0 && (6.76)\\
\frac{\partial \dot{\boldsymbol{p}}_I}{\partial \boldsymbol{p}_I} &= -D\,N_I\,(\xi_I + \xi_P) && (6.77)\\
\frac{\partial \dot{\boldsymbol{p}}_{II}}{\partial \boldsymbol{p}_{II}} &= -D\,N_{II}\,\xi_{II} && (6.78)\\
\frac{\partial \dot{\boldsymbol{p}}_{f,I}}{\partial \boldsymbol{p}_{f,I}} &= -D\,\nu_I\,(\xi_I + \xi_P) && (6.79)\\
\frac{\partial \dot{\boldsymbol{p}}_{f,II}}{\partial \boldsymbol{p}_{f,II}} &= -D\,\nu_{II}\,\xi_{II} && (6.80)\\
\frac{\partial \dot{\xi}_I}{\partial \xi_I} &= 0 && (6.81)\\
\frac{\partial \dot{\xi}_{II}}{\partial \xi_{II}} &= 0 && (6.82)\\
\frac{\partial \dot{\eta}_I}{\partial \eta_I} &= 0 && (6.83)\\
\frac{\partial \dot{\eta}_{II}}{\partial \eta_{II}} &= 0 && (6.84)\\
\frac{\partial \dot{V}_I}{\partial V_I} &= D\,\xi_P && (6.85)\\
\frac{\partial \dot{\xi}_P}{\partial \xi_P} &= 0 && (6.86)\\
\frac{\partial \dot{\nu}_I}{\partial \nu_I} &= 0 && (6.87)\\
\frac{\partial \dot{p}_{\nu_I}}{\partial p_{\nu_I}} &= 0 && (6.88)\\
\kappa(\boldsymbol{\eta}) &\equiv -D\,(N_I + \nu_I - 1)\,\xi_P - D\,(N_I + \nu_I)\,\xi_I - D\,(N_{II} + \nu_{II})\,\xi_{II} && (6.89)
\end{aligned}$$

The metric determinant factor $\sqrt{g}$, treating ν_I and ν_{II} as constant, is:

$$\begin{aligned}
\kappa(\boldsymbol{\eta}) &= \dot{w}(\boldsymbol{\eta}) && (6.90)\\
\dot{w}(\boldsymbol{\eta}) &= -D\,(N_I+\nu_I-1)\,\xi_P - D\,(N_I+\nu_I)\,\xi_I - D\,(N_{II}+\nu_{II})\,\xi_{II} && (6.91)\\
w(\boldsymbol{\eta}) &= -D\,(N_I+\nu_I-1)\,\epsilon - D\,(N_I+\nu_I)\,\eta_I - D\,(N_{II}+\nu_{II})\,\eta_{II} && (6.92)\\
\sqrt{g(\boldsymbol{\eta})} &= \mathrm{e}^{-w(\boldsymbol{\eta})} && (6.93)\\
&= \mathrm{e}^{D\,(N_I+\nu_I-1)\,\epsilon}\,\mathrm{e}^{D\,(N_I+\nu_I)\,\eta_I}\,\mathrm{e}^{D\,(N_{II}+\nu_{II})\,\eta_{II}} && (6.94)
\end{aligned}$$

The conserved energy is:

$$\begin{aligned}
\mathcal{H}' &= \sum_{i=1}^{N_I}\frac{\boldsymbol{p}_{I,i}^2}{2\,m_{I,i}} + \Phi(\boldsymbol{q}_I)\\
&\quad + \sum_{i=1}^{N_{II}}\frac{\boldsymbol{p}_{II,i}^2}{2\,m_{II,i}} + \Phi(\boldsymbol{q}_{II})\\
&\quad + \frac{1}{2}\,Q_I\,\xi_I^2 + g_I\,k_B\,T_E\,\eta_I\\
&\quad + \frac{1}{2}\,Q_{II}\,\xi_{II}^2 + g_{II}\,k_B\,T_E\,\eta_{II}\\
&\quad + \frac{1}{2}\,Q_P\,\xi_P^2\\
&\quad + \mathcal{H}_{\mathrm{xch}} \qquad (6.95)
\end{aligned}$$

The partition function can now be calculated knowing the metric determinant factor and the conserved energy.

$$\begin{aligned}
\Omega_{\mathrm{GEMD}} &= \int \mathrm{d}^{N_I}\boldsymbol{p}_I \int_0^1 \mathrm{d}^{N_I}\boldsymbol{r}_I \int \mathrm{d}^{\nu_I}\boldsymbol{p}_{f,I} \int_0^1 \mathrm{d}^{\nu_I}\boldsymbol{r}_{f,I} \int \mathrm{d}\xi_I \int \mathrm{d}\eta_I\\
&\quad \times \int \mathrm{d}^{N_{II}}\boldsymbol{p}_{II} \int_0^1 \mathrm{d}^{N_{II}}\boldsymbol{r}_{II} \int \mathrm{d}^{\nu_{II}}\boldsymbol{p}_{f,II} \int_0^1 \mathrm{d}^{\nu_{II}}\boldsymbol{r}_{f,II} \int \mathrm{d}\xi_{II} \int \mathrm{d}\eta_{II}\\
&\quad \times \int_0^V \mathrm{d}V_I \int \mathrm{d}\xi_P \int_0^1 \mathrm{d}\nu_I \int \mathrm{d}p_{\nu_I}\,\sqrt{g}\\
&\quad \times \delta\Big[\mathcal{H}_0\left(\boldsymbol{p}_I, \boldsymbol{r}_I\,V_I^{\frac{1}{D}}\right) + \frac{1}{2}\,Q_I\,\xi_I^2 + g_I\,k_B\,T_E\,\eta_I + \frac{1}{2}\,Q_P\,\xi_P^2\\
&\quad + \mathcal{H}_0\left(\boldsymbol{p}_{II}, \boldsymbol{r}_{II}\,V_{II}^{\frac{1}{D}}\right) + \frac{1}{2}\,Q_{II}\,\xi_{II}^2 + g_{II}\,k_B\,T_E\,\eta_{II} + \mathcal{H}_{\mathrm{xch}} - E\Big] \qquad (6.96)
\end{aligned}$$

An integration over η_I is done, leading to the final partition function generated by the Gibbs ensemble molecular dynamics equations of motion.

$$\begin{aligned}
\eta_{I,0} &= \frac{E - \mathcal{H}_{\text{xch}} - \mathcal{H}_0\left(\boldsymbol{p}_I, \boldsymbol{r}_I\, V_I^{\frac{1}{D}}\right) - \frac{1}{2}\, Q_I\, \xi_I^2 - \frac{1}{2}\, Q_P\, \xi_P^2}{g_I\, k_B\, T_E} \\
&\quad + \frac{-\mathcal{H}_0\left(\boldsymbol{p}_{II}, \boldsymbol{r}_{II}\, V_{II}^{\frac{1}{D}}\right) - \frac{1}{2}\, Q_{II}\, \xi_{II}^2 - g_{II}\, k_B\, T_E\, \eta_{II}}{g_I\, k_B\, T_E} \qquad (6.97) \\
f'(\eta_I) &= g_I\, k_B\, T_E \qquad (6.98) \\
\epsilon &= \frac{1}{D}\, \ln \frac{V_I}{V_0} \qquad (6.99) \\
\Omega_{\text{GEMD}} &= \frac{1}{g_I\, k_B\, T_E}\, \exp\left(\frac{D\,(N_I + \nu_I)}{g_I}\, \beta\, E\right) \\
&\quad \times \int \mathrm{d}^{N_I}\boldsymbol{p}_I \int_0^1 \mathrm{d}^{N_I}\boldsymbol{r}_I \exp\left(-\frac{D\,(N_I + \nu_I)}{g_I}\, \beta\, \mathcal{H}_0\left(\boldsymbol{p}_I, \boldsymbol{r}_I\, V_I^{\frac{1}{D}}\right)\right) \\
&\quad \times \int \mathrm{d}^{N_{II}}\boldsymbol{p}_{II} \int_0^1 \mathrm{d}^{N_{II}}\boldsymbol{r}_{II} \exp\left(-\frac{D\,(N_I + \nu_I)}{g_I}\, \beta\, \mathcal{H}_0\left(\boldsymbol{p}_{II}, \boldsymbol{r}_{II}\, V_{II}^{\frac{1}{D}}\right)\right) \\
&\quad \times \int \mathrm{d}^{\nu_I}\boldsymbol{p}_{f,I} \int_0^1 \mathrm{d}^{\nu_I}\boldsymbol{r}_{f,I} \int \mathrm{d}^{\nu_{II}}\boldsymbol{p}_{f,II} \int_0^1 \mathrm{d}^{\nu_{II}}\boldsymbol{r}_{f,II} \int_0^1 \mathrm{d}\nu_I \int \mathrm{d}p_{\nu_I} \\
&\quad \times \exp\left(-\frac{D\,(N_I + \nu_I)}{g_I}\, \beta\, \mathcal{H}_{\text{xch}}\right) \\
&\quad \times \int \mathrm{d}\xi_I\, \exp\left(-\frac{D\,(N_I + \nu_I)}{g_I}\, \frac{\beta}{2}\, Q_I\, \xi_I^2\right) \int \mathrm{d}\xi_{II}\, \exp\left(-\frac{D\,(N_I + \nu_I)}{g_I}\, \frac{\beta}{2}\, Q_{II}\, \xi_{II}^2\right) \\
&\quad \times \int \mathrm{d}\eta_{II}\, \exp\left(D\,(N_{II} + \nu_{II})\, \eta_{II}\right) \exp\left(-\frac{D\,(N_I + \nu_I)}{g_I}\, g_{II}\, \eta_{II}\right) \\
&\quad \times \int_0^V \mathrm{d}V_I\, \left(\frac{V_I}{V_0}\right)^{N_I + \nu_I - 1} \\
&\quad \times \int \mathrm{d}\xi_P\, \exp\left(-\frac{D\,(N_I + \nu_I)}{g_I}\, \frac{\beta}{2}\, Q_P\, \xi_P^2\right) \qquad (6.100)
\end{aligned}$$

Setting g_I equal to $D\,(N_I + \nu_I)$ and g_{II} to $D\,(N_{II} + \nu_{II})$ gives a zero in the exponent of the η_{II} integral. The solution of the general integral is infinite. This situation is comparable to the divergence of the s_2' integral seen by Palmer and Lo. The integral diverges for all values of g_{II}. In reality, the energy transferred from or to the heat bath is limited by the conserved ensemble energy. The integral has an upper and lower limit. Therefore the integral solution is a constant value and never infinite. If expectation values are calculated then both identical integrals in the denominator and numerator will be canceled.

The total degrees of freedom g and the number of particles are:

$$g = D\,N \tag{6.101}$$
$$N = N_I + N_{II} + \nu_I + \nu_{II} \tag{6.102}$$

Considering now the special case without a fractional particle. Both ν_I and ν_{II} will be zero and the exchange Hamiltonian drops out.

The barostat integrals give a constant factor. The partition function is proportional to the canonical partition function.

$$\begin{aligned}\Omega_{\mathrm{GEMD}} \propto & \int \mathrm{d}^{N_I}\boldsymbol{p}_I \int_0^1 \mathrm{d}^{N_I}\boldsymbol{r}_I \, \exp\left(-\beta\,\mathcal{H}_0\left(\boldsymbol{p}_I, \boldsymbol{r}_I\, V_I^{\frac{1}{D}}\right)\right) \\ & \times \int \mathrm{d}^{N_{II}}\boldsymbol{p}_{II} \int_0^1 \mathrm{d}^{N_{II}}\boldsymbol{r}_{II} \, \exp\left(-\beta\,\mathcal{H}_0\left(\boldsymbol{p}_{II}, \boldsymbol{r}_{II}\, V_{II}^{\frac{1}{D}}\right)\right)\end{aligned} \tag{6.103}$$

The molecular dynamics part produces the same ($\mathrm{N_I}$+$\mathrm{N_{II}}$)($\mathrm{V_I}$+$\mathrm{V_{II}}$)T ensemble as in the Gibbs ensemble Monte Carlo method, if no particle transfers are done.

Exchange Probability

The probability of transferring a particle from system II to system I is given by the ratio of two phase space probability density functions. Both probability density functions are multiplied by the compressibility of the first thermostat and an integration over η_I is done. Obtaining a ratio of two partition functions.

$$\frac{\wp(\nu_I = 1)}{\wp(\nu_I = 0)} = \frac{f(\nu_I = 1)}{f(\nu_I = 0)} \tag{6.104}$$
$$= \frac{\int \mathrm{d}\eta_I \, \mathrm{e}^{g_I\,\eta_I}\, f(\nu_I = 1)}{\int \mathrm{d}\eta_I \, \mathrm{e}^{g_I\,\eta_I}\, f(\nu_I = 0)} \tag{6.105}$$
$$= \frac{(N_{II}+1)\,V_I}{(N_I+1)\,V_{II}}\, \mathrm{e}^{-\beta\,(\mathcal{H}_I(N_I+1)-\mathcal{H}_I(N_I)-(\mathcal{H}_{II}(N_{II}+1)-\mathcal{H}_{II}(N_{II})))} \tag{6.106}$$
$$= \frac{(N_{II}+1)\,V_I}{(N_I+1)\,V_{II}}\, \mathrm{e}^{-\beta\,\Delta\,E_{\mathrm{xch}}} \tag{6.107}$$
$$= \frac{\rho_{II}}{\rho_I}\, \frac{\mathrm{e}^{-\beta\,(\mathcal{H}_I(N_I+1)-\mathcal{H}_I(N_I))}}{\mathrm{e}^{-\beta\,(\mathcal{H}_{II}(N_{II}+1)-\mathcal{H}_{II}(N_{II}))}} \tag{6.108}$$

The formally correct transfer probability is obtained which satisfies the equilibrium condition that the chemical potential is equal in both phases (see below). The Monte Carlo probability to create a particle in region I is given in equation (6.5) on page 54.

The chemical potential of a component in the NVT ensemble is defined as:

$$\mu = \left(\frac{\partial A}{\partial n}\right)_{V,T,n_{i\neq j}} \tag{6.109}$$

The Helmholtz free energy is given in equation (2.79) on page 15. The chemical potential in both phases is calculated by the two partition functions in equation (6.108).

$$\mu_I = k_B T_E \ln \rho_I - k_B T_E \ln \left(\mathrm{e}^{-\beta\,(\mathcal{H}_I(N_I+1)-\mathcal{H}_I(N_I))}\right) \tag{6.110}$$

$$\mu_{II} = k_B T_E \ln \rho_{II} - k_B T_E \ln \left(\mathrm{e}^{-\beta\,(\mathcal{H}_{II}(N_{II}+1)-\mathcal{H}_{II}(N_{II}))}\right) \tag{6.111}$$

Which is equal to Widom's equation for a test particle[95]:

$$\mu = k_B T \ln \rho - k_B T \ln \left\langle \mathrm{e}^{-\beta\,\Delta U^+} \right\rangle \tag{6.112}$$

Calculating the chemical potential of the entire ensemble by equation (6.108) gives the difference of the chemical potentials in both phases:

$$\mu = \mu_I - \mu_{II} \tag{6.113}$$

Which is zero, if the phases are in equilibrium.

6.3. Other Algorithms

An alternative molecular dynamics algorithm, in which the chemical potential is not known was published by Kotelyanskii and Hentschke[43]. In their algorithm, all particles have an additional degree of freedom, which determines the particles' position between the subsystems. Results were published for Lennard-Jones systems and for hexane[33]. The hexane results are so far the only published molecular results. Palmer and Lo did not publish any molecular data.

7. Gibbs Ensemble Molecular Dynamics Simulations

The Gibbs ensemble molecular dynamics method as described in the previous chapter was implemented and used to calculate the critical properties of *ab initio* CO_2 potentials. The definition of the potentials, the parameterization, the obtained equilibrium values and finally the fit of the critical values are described in this chapter.

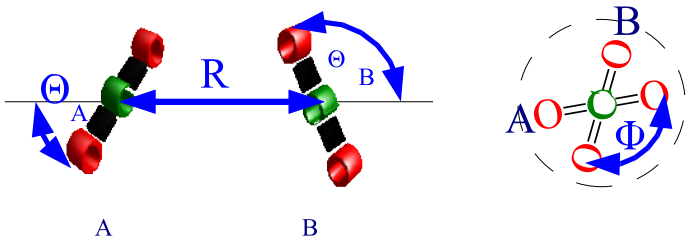
The common parameters of all simulations are defined in table 7.1.

Parameter	**Value**
Molecular model:	Rigid molecules.
Initial geometry:	Cubic lattice with random orientations.
Number of molecules:	250 molecules.
Time step:	1 fs.
ν_0:	0
Exchange start:	After 100 time steps.
Max. exchange break:	100 time steps.
Histogram:	Every 100th value used.

Table 7.1.: Common Gibbs Ensemble Simulation Parameters.

The experimental reference values are from reference [6].

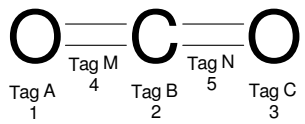
7.1. CO_2 Potentials



The potential energy of the two rigid molecules is a function of four variables. The distance R,

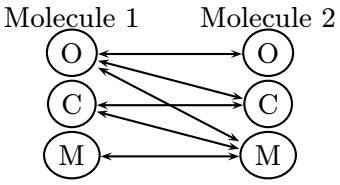
the angles Θ_A and Θ_B between the distance R and the axis of the molecule A and B, and the dihedral angle Φ.

Site-site potential functions are acquired from fits to this four dimensional hypersurface.



$1 \equiv 3$ $\qquad\qquad$ $4 \equiv 5$

Dimer interactions:



6 unique site-site interactions are given.

Two additional sites, without a mass, are placed on the CO bonds to improve the fit and charge distribution.

7.1.1. The BBV Potential

The Bock-Bich-Vogel (BBV) potential function[10] has 8 parameters per site-site interaction. It uses the potential form developed by Steinebrunner *et al.*[77] for the 8s6p4d1f potential.

$$d(r_{ij}) = \left(1 + \mathrm{e}^{-C_7\, r_{ij} + C_8}\right)^{-\gamma} \tag{7.1}$$

$$\Phi(r_{ij}) = C_5\, \mathrm{e}^{-C_6\, r_{ij}} + \left(\frac{C_1}{r_{ij}^{12}} + \frac{C_2}{r_{ij}^{10}} + \frac{C_3}{r_{ij}^{8}} + \frac{C_4}{r_{ij}^{6}} + \frac{q_i q_j}{r_{ij}}\right) d(r_{ij}) \tag{7.2}$$

Where equation (7.1) defines the damping function for the polynomial terms. The constant γ has the value 15.

Energies for 1060 unique dimer configurations were calculated at MP2 level with full counter-poise correction. The basis sets cc-pVDZ, cc-pVTZ and cc-pVQZ from the correlation consistent series of Dunning[19] where used and the final points for the site-site fitting were obtained from an exponential extrapolation to a complete basis-set.

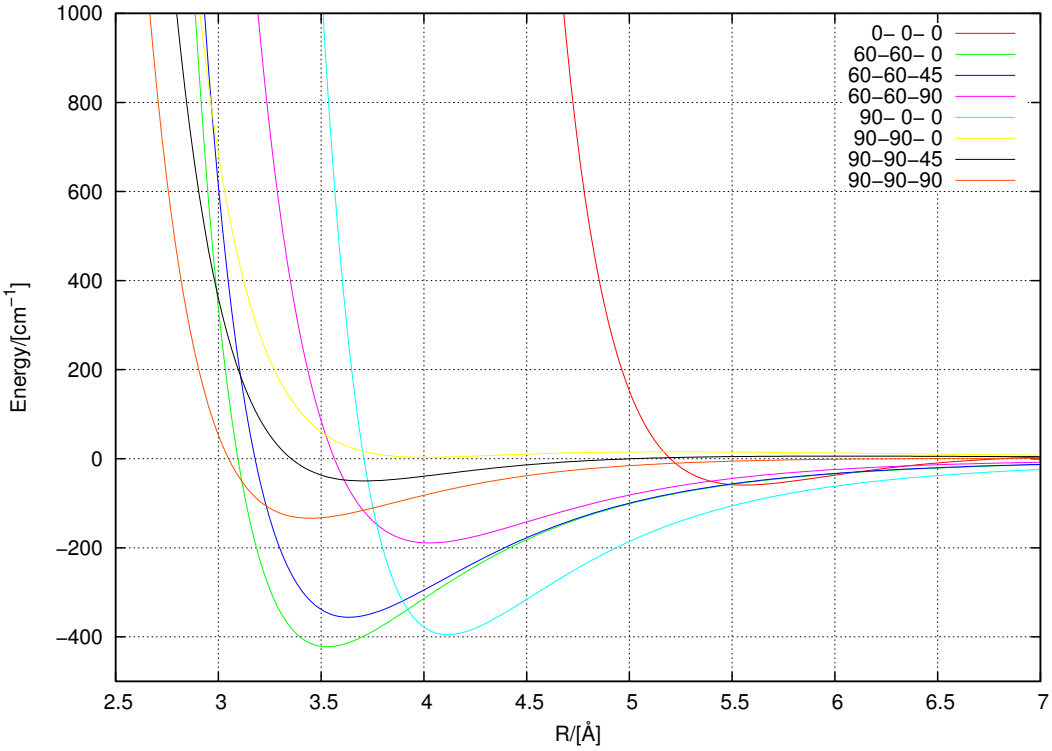


Figure 7.1.: Cuts Through the BBV Dimer Potential.

Θ_A	Θ_B	Φ	**Name**
0	0	0	Linear (repulsive)
60	60	0	Slipped parallel (minimum)
90	0	0	T-shaped
90	90	0	Parallel
90	90	90	Crossed

The CO bond length is 1.161 Å, the oxygen charge -0.3321 $|e|$ and the carbon charge 0.6642 $|e|$. The bond length is in the range of experimental values, the charges reproduce the experimental quadrupole moment of the monomer. The dummy sites are in the middle of the CO bond and are not charged. The potential values are pure *ab initio*, no scaling to experimental values was done.

Intramolecular Potential

An intramolecular potential for a completely flexible molecule was calculated by Vogt[92] with Gaussian 98[25]. The frequencies obtained by MP2 calculations with the basis sets cc-pVDZ, cc-pVTZ and cc-pVQZ were extrapolated to a complete basis. The resulting force constants and the equilibrium bond length are:

$r_e/[\AA]$	$k_\alpha/[N\,rad]$	$k_{CO}/[N\,m]$	$k_{OO}/[N\,m]$
1.164	$0.71899\text{x}10^{-18}$	1506.6	86.515

Flexible molecules were used for the local density calculations (see chapter 8 on page 105). Comparisons to rigid models showed that comparable results were obtained. A limiting factor of flexible molecules is the time step, 0.25 fs is the optimum value, whereas for rigid molecules this value could be increased to 1 fs in liquid simulations. A factor of 2-3 for the time step was used by other authors[44]. The time step can be increased because the fastest motion in the system is no longer the intramolecular vibration. In the gas phase, larger values are possible than in a dense liquid phase, where the translational motion comes closer to a vibration.

Another problem of treating the intramolecular motion by a classical harmonic potential is the negligence of quantum effects at higher temperatures. It was shown[82] that using a rigid bond length is often a better approximation than a classical vibration. Tironi, Brunne and van Gunsteren write:

> The introduction of flexibility is shown to create more problems than it solves and to lead to less accurate results in the case of liquid water simulations at room temperature and pressure.

Pressure Comparison

One important check of the quality of a potential is how well it reproduces the experimental pressure in a high density state, the liquid phase.

The values in the table below are obtained from molecular dynamics simulations of 343 rigid CO_2 molecules with 20000 equilibration steps and 20000 production steps in the canonical ensemble. A Nosé-Hoover thermostat was used to control the temperature. The external temperature was 280 K. The fundamental time step was 1 fs. The standard deviation of the temperature is below 0.1 K and not shown here.

Density$/[\frac{mol}{m^3}]$	**8s6p4d1f[77]**$/[bar]$	**BBV**$/[bar]$	**Exp.[72]**$/[bar]$
20590	270	88 ± 1	60
21818	300	130 ± 1	130
22695	450	208 ± 1	210
23389	570	284 ± 1	290
25665	1080	677 ± 1	720

Table 7.2.: CO_2 280 K Pressure Comparison

The BBV values show excellent agreement with the experimental data. The improvement from the 8s6p4d1f potential is obvious.

The larger differences at higher densities can be explained by the missing manybody interactions. The 20590 $\mathrm{mol\,m^{-3}}$ value is already close to the two-phase region. Gibbs ensemble simulations showed that the two-phase region starts at 20120 $\mathrm{mol\,m^{-3}}$. Therefore the simulation already samples some interface effects, giving unreliable results for the pressure.

7.1.2. The SAPT-s Potential

The SAPT-s potential function[12] has 7 parameters per site-site interaction.

$$\Phi(r_{ij}) = \mathrm{e}^{\alpha_{ij}-\beta_{ij}r_{ij}} + f_1\left(\delta_1^{ij} r_{ij}\right)\frac{q_i q_j}{r_{ij}} - f_6\left(\delta_6^{ij} r_{ij}\right)\frac{C_6^{ij}}{r_{ij}^6} - f_8\left(\delta_8^{ij} r_{ij}\right)\frac{C_8^{ij}}{r_{ij}^8} \tag{7.3}$$

The damping functions of Tang-Toennies type[81] are defined as:

$$f_n[x] = 1 - \mathrm{e}^{-x}\sum_{k=0}^{n}\frac{x^k}{k!} \tag{7.4}$$

$$f_1 = 1 - \mathrm{e}^{-x}(1+x) \tag{7.5}$$

$$f_6 = 1 - \mathrm{e}^{-x}\left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720}\right) \tag{7.6}$$

$$f_8 = 1 - \mathrm{e}^{-x}\left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720} + \frac{x^7}{5040} + \frac{x^8}{40320}\right) \tag{7.7}$$

220 dimers were calculated with many-body symmetry-adapted perturbation theory with a 5s3p2d1f basis set including bond functions.

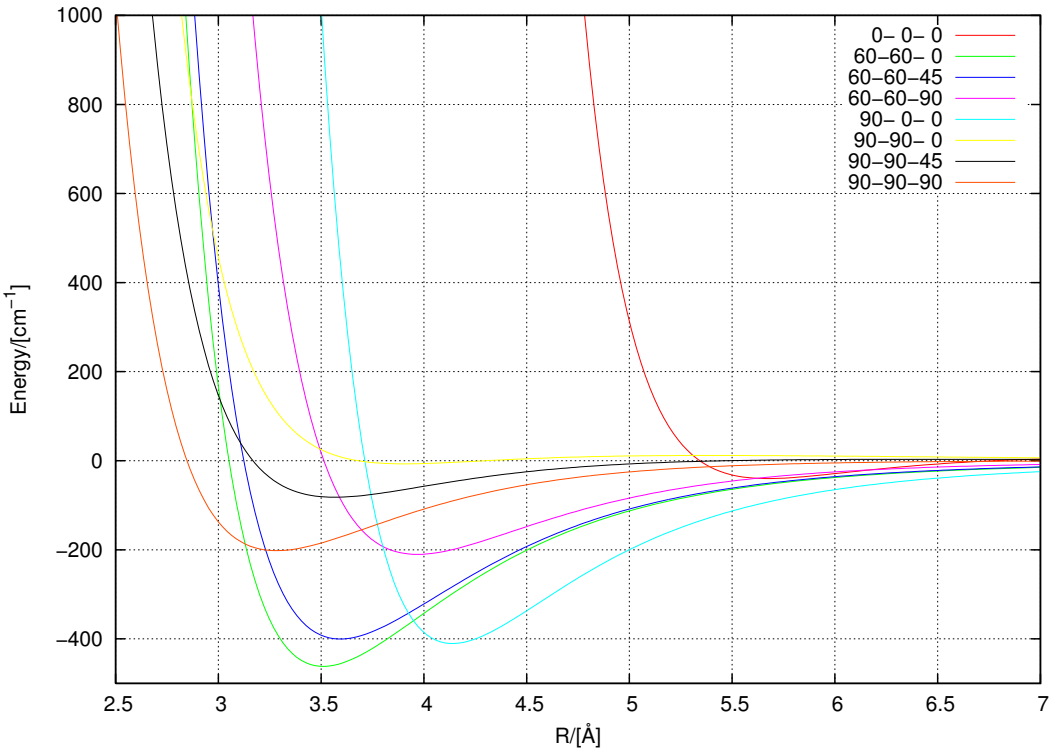


Figure 7.2.: Cuts Through the SAPT Dimer Potential.

Each molecule has five charge sites. The CM distance (carbon-dummy site) is 0.8456 Å and the CO distance 1.162047 Å. The dummy sites are closer to the oxygen. The bond length was obtained from the experimental rotational constant B_0 where the isotropically averaged quantities were used.

The elementary charges are:

Atom	**Charge**/[$\|e\|$]
q_O	0.23786535
q_C	1.6316722
q_M	-1.0537015

All atoms are positively charged and the negative charge is placed onto the dummy sites close to the oxygen. The fastest polynomial decay of the long-range part of the potential is proportional to r^{-8} where the BBV potential has in addition r^{-10} and r^{-12} terms. The long-range corrections of the pressure are mostly negligible for the BBV potential, the SAPT-s values are approximately 100 times higher and no longer negligible, even for extremely large cut-off values.

7.2. Single Phase Region

To check the stability of the simulation program, several runs were done in the single phase region. The parameters used for the simulations are shown in table 7.3.

Parameter	**Value**
Potential:	BBV
Temperature:	280 K
Number of molecules:	250 molecules.
Density I:	25665 mol m^{-3}
Density II:	25665.01 mol m^{-3}
Q:	1x10^{-44} kg m^2
Q_P:	1x10^{-42} kg m^2
Q_E:	1x10^{-20} kg m^2
η	4 Å
v_0	0.001

Table 7.3.: BBV 280 K Liquid: Single Phase Simulation Parameters.

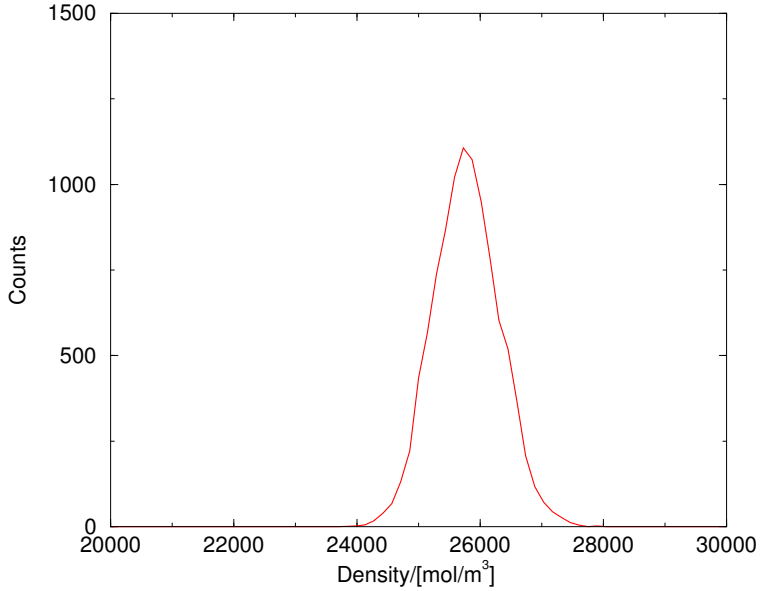


Figure 7.3.: BBV 280 K Liquid: Density Histogram of Both Subsystems.

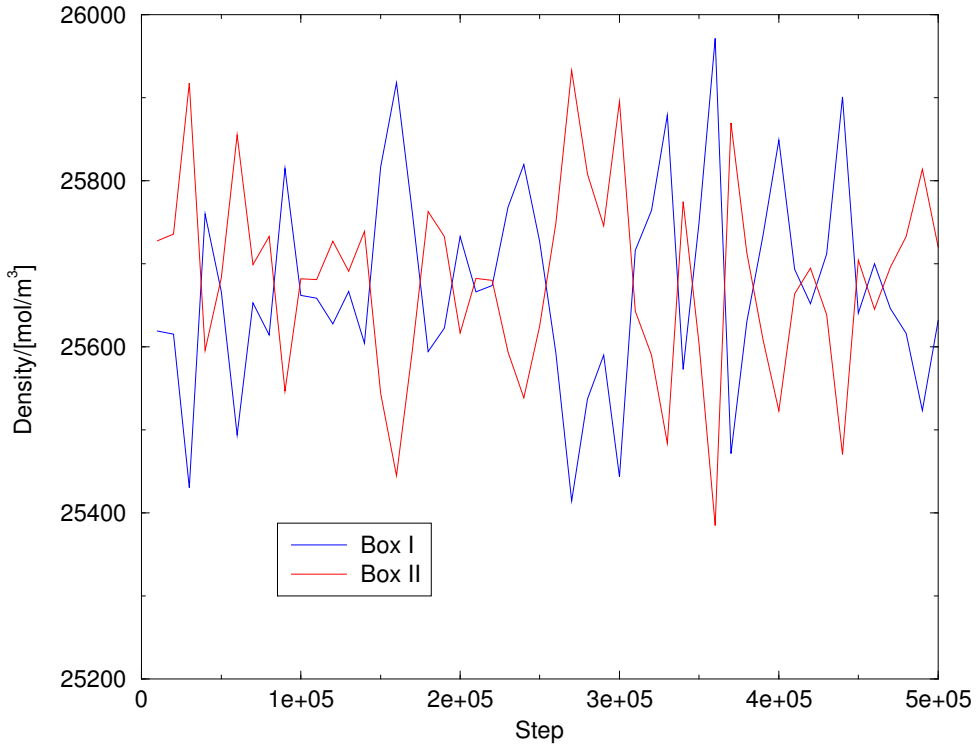


Figure 7.4.: BBV 280 K Liquid: Densities Averaged over 10000 Steps.

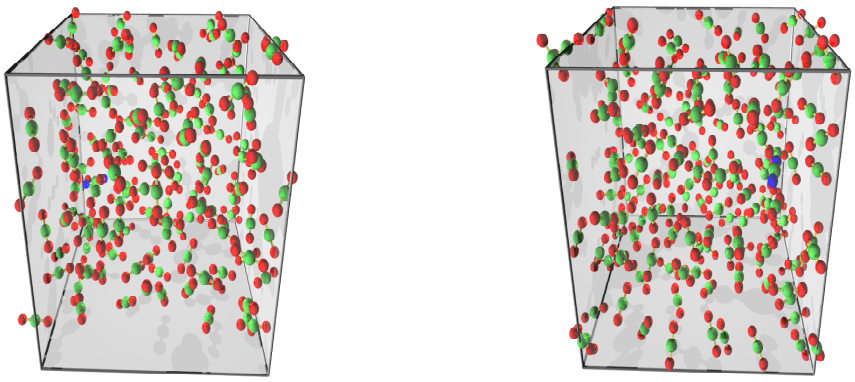


Figure 7.5.: BBV 280 K Liquid: Single Phase Snapshot after Successful Exchange.

The snapshot of both simulation boxes in figure 7.5 shows on the left the source box and on the right the destination box of the particle transfer. The fractional molecule was successfully integrated into the destination system. The oxygens of the fractional molecule are colored in blue.

Graphs representing the volume and number of molecules are not shown here, because both subsystems are in equilibrium and particle exchanges in both directions are possible without changing the density. If one subsystem looses particles, the volume is automatically decreased to maintain the pressure.

The final results of 500,000 time steps are shown in table 7.4.

Parameter	**Value**
Temperature I:	279.96 K
Temperature II:	280.03 K
Density I:	25673 $\mathrm{mol\,m^{-3}}$
Density II:	25679 $\mathrm{mol\,m^{-3}}$
Pressure I:	646 bar
Pressure II:	641 bar
Average pressure:	643 bar
Exchanges:	40

Table 7.4.: BBV 280 K Liquid: Single Phase Results.

The temperature is controlled by the Nosé-Hoover thermostat. Both values are close to the external temperature. The average temperature of both subsystems is even closer, generating as a whole the expected NVT ensemble. The pressure is not fixed to an external value but both subsystems are coupled to have the same pressure. The difference to the average pressure is negligible showing that the barostatization is efficient. The density histogram 7.3 shows a Gaussian distribution of the densities. The quality of the histogram is only limited by the 50 $\mathrm{mol\,m^{-3}}$ resolution.

7.3. BBV Results

Simulations were done in steps of 10 K from 220 K to 300 K. The triple point of CO_2 is at 217 K and 5.2 bar. The potential cut-off radius was always larger than 10 Å, where the g-function is equal to the distribution of an ideal gas (see figure 8.2 on page 108). The averages were taken over 500,000 steps of continued simulations, giving a total simulation length of several million time steps. The common parameters, if not otherwise defined, for each run were:

Parameter	**Value**
Number of molecules:	250 molecules.
Time step:	1 fs
Equilibration:	40,000 steps using velocity rescaling
Production:	500,000 steps per run
Q:	$1\text{x}10^{-44}\,\text{kg}\,\text{m}^2$
Q_P:	$1\text{x}10^{-42}\,\text{kg}\,\text{m}^2$
Q_E:	$1\text{x}10^{-21}\,\text{kg}\,\text{m}^2$
ν_0:	0
v_0	0.001
η	4 Å
Exchange start:	After 100 time steps.
Max. exchange break:	100 time steps.
Histogram:	Every 100th value used.

Table 7.5.: BBV Gibbs Ensemble Molecular Dynamics Parameters.

The exchange mass Q_E used in the single phase region (see table 7.3) was optimized for the liquid-vapor simulations.

7.3.1. 220 K

The barostat mass was reduced to $1\text{x}10^{-41}\,\text{kg}\,\text{m}^2$ to better control the fluctuations between both phases. At 220 K the volume difference is high and the density of the gas phase is already close to zero.

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
434	26851	7.1
410	26904	6.7
428	26743	6.9
410	26906	6.8

Table 7.6.: BBV 220 K: Results.

Property	**Value**	
ρ_g:	421 ± 6	$\frac{mol}{m^3}$
ρ_l:	26851 ± 38	$\frac{mol}{m^3}$
P:	6.9 ± 0.1	bar
$\rho_{g,\mathrm{exp}}$:	364.3	$\frac{mol}{m^3}$
$\rho_{l,\mathrm{exp}}$:	26499	$\frac{mol}{m^3}$
P_{exp}:	5.996	bar

Table 7.7.: BBV 220 K: Averages.

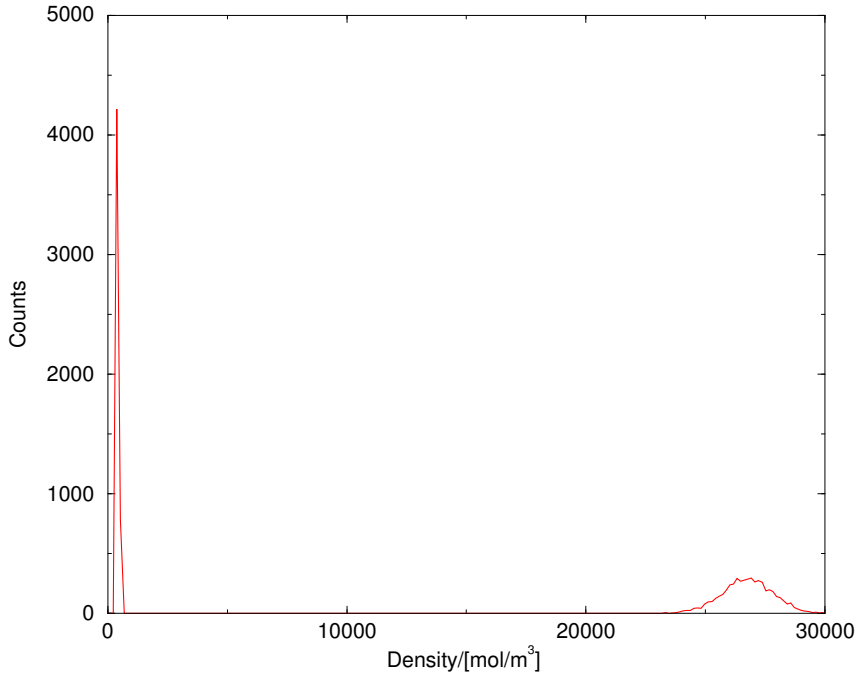


Figure 7.6.: BBV 220 K: Density Histogram.

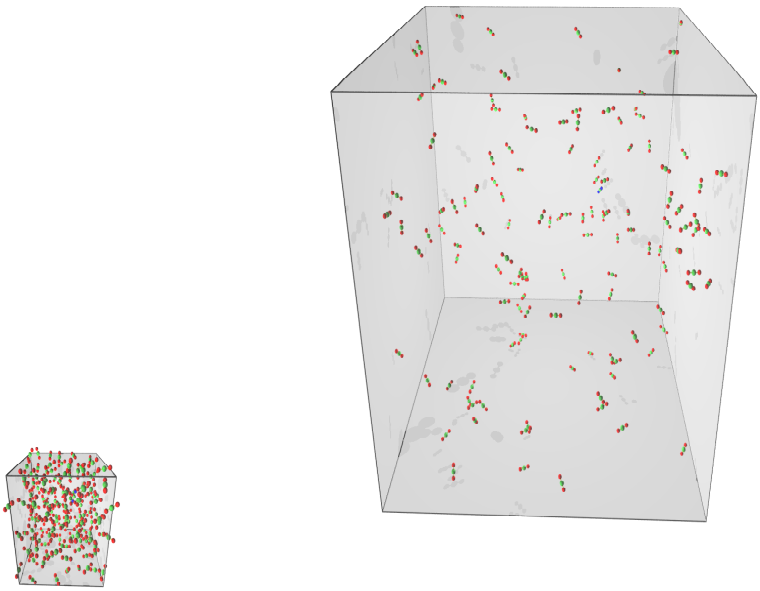


Figure 7.7.: BBV 220 K Snapshot.

7.3.2. 230 K

The barostat mass was reduced to $1\text{x}10^{-41}\,\text{kg}\,\text{m}^2$.

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
585	25990	9.8
587	25957	9.9
604	25939	10.2
581	26057	9.6
580	25828	9.6
594	25855	9.7
603	25978	9.9
603	25786	9.8

Table 7.8.: BBV 230 K: Results.

Property	Value	
ρ_g:	592 ± 4	$\frac{mol}{m^3}$
ρ_l:	25924 ± 33	$\frac{mol}{m^3}$
P:	9.8 ± 0.1	bar
$\rho_{g,\mathrm{exp}}$:	531.1	$\frac{mol}{m^3}$
$\rho_{l,\mathrm{exp}}$:	25655	$\frac{mol}{m^3}$
P_{exp}:	8.935	bar

Table 7.9.: BBV 230 K: Averages.

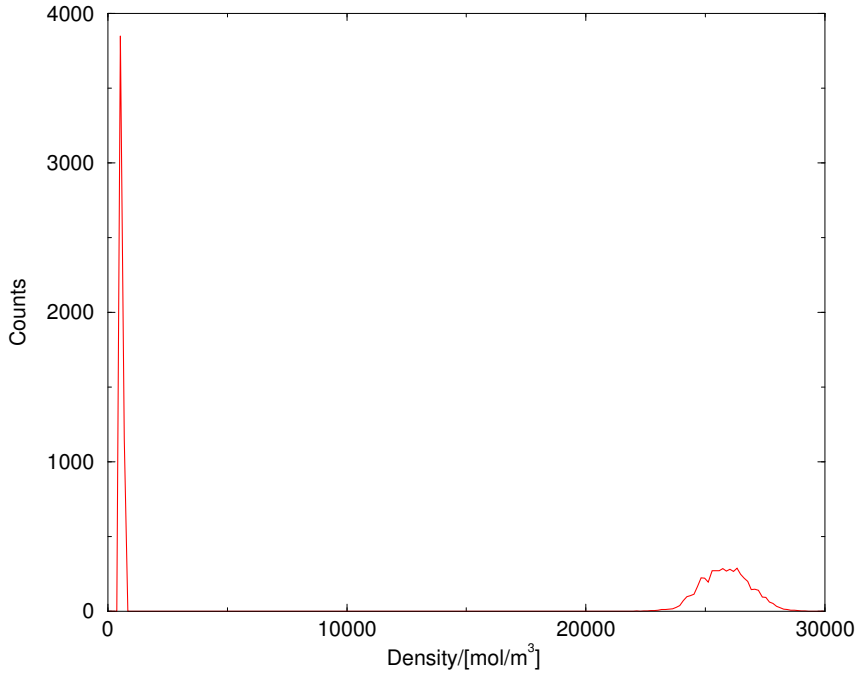


Figure 7.8.: BBV 230 K: Density Histogram.

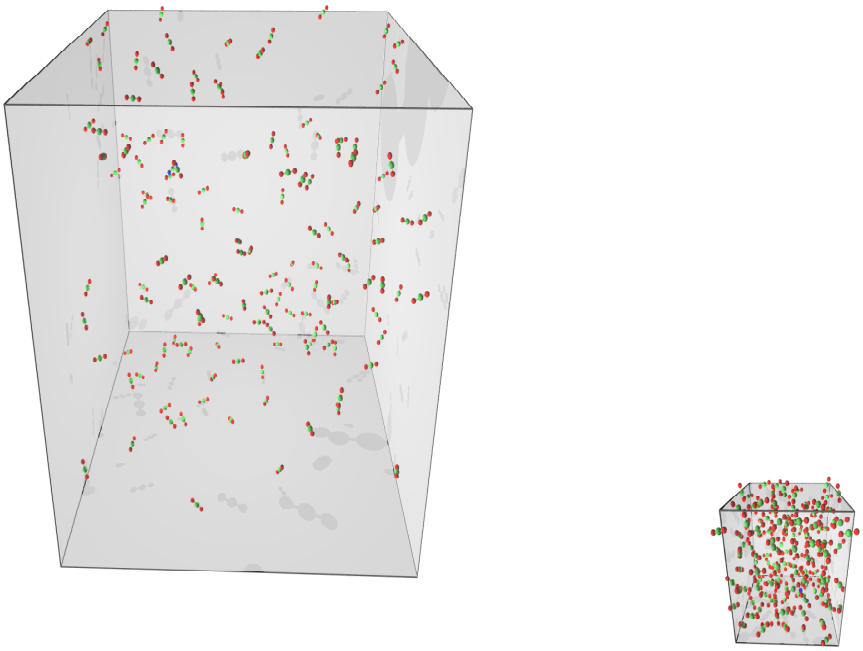


Figure 7.9.: BBV 230 K Snapshot.

7.3.3. 240 K

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
816	24984	13.7
865	24989	14.4
860	24927	14.5
874	25028	14.9
837	24976	14.1
815	24957	13.8
885	24990	14.8

Table 7.10.: BBV 240 K: Results.

Property	**Value**	
ρ_g:	850 ± 11	$\frac{mol}{m^3}$
ρ_l:	24979 ± 12	$\frac{mol}{m^3}$
P:	14.3 ± 0.2	bar
$\rho_{g,\mathrm{exp}}$:	758.3	$\frac{mol}{m^3}$
$\rho_{l,\mathrm{exp}}$:	24754	$\frac{mol}{m^3}$
P_{exp}:	12.830	bar

Table 7.11.: BBV 240 K: Averages.

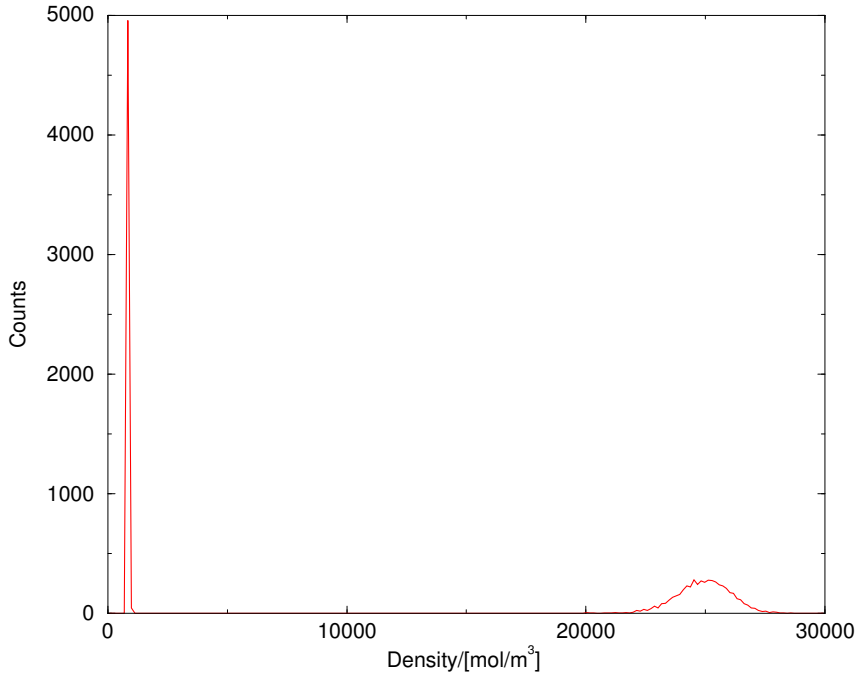


Figure 7.10.: BBV 240 K: Density Histogram.

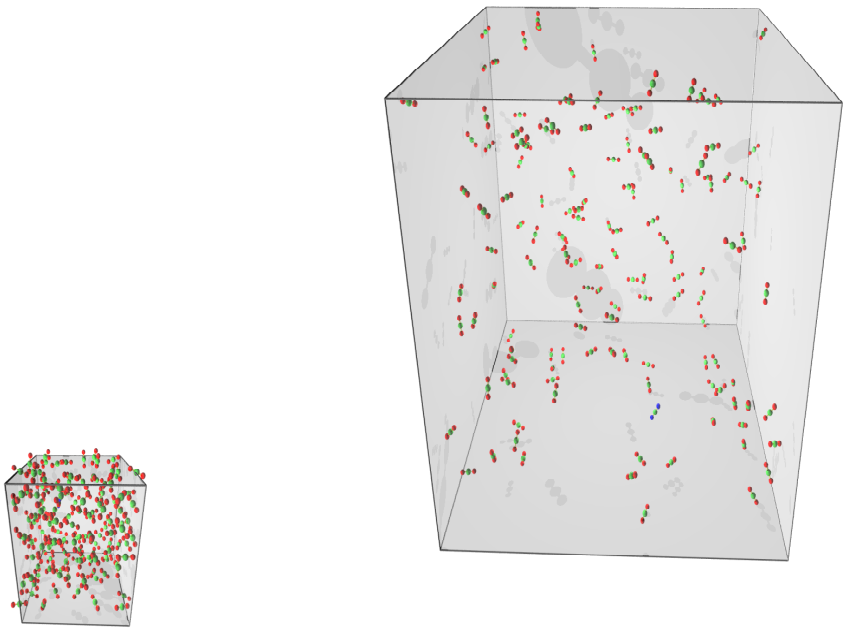


Figure 7.11.: BBV 240 K Snapshot.

7.3.4. 250 K

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
1438	23961	22.9
1560	23711	24.0
1523	23687	23.8
1474	23858	22.6
1517	23571	23.9

Table 7.12.: BBV 250 K: Results.

Property	**Value**
ρ_g:	1502 ± 21 $\frac{mol}{m^3}$
ρ_l:	23758 ± 68 $\frac{mol}{m^3}$
P:	23.4 ± 0.3 bar
$\rho_{g,\mathrm{exp}}$:	1061.6 $\frac{mol}{m^3}$
$\rho_{l,\mathrm{exp}}$:	23781 $\frac{mol}{m^3}$
P_{exp}:	17.856 bar

Table 7.13.: BBV 250 K: Averages.

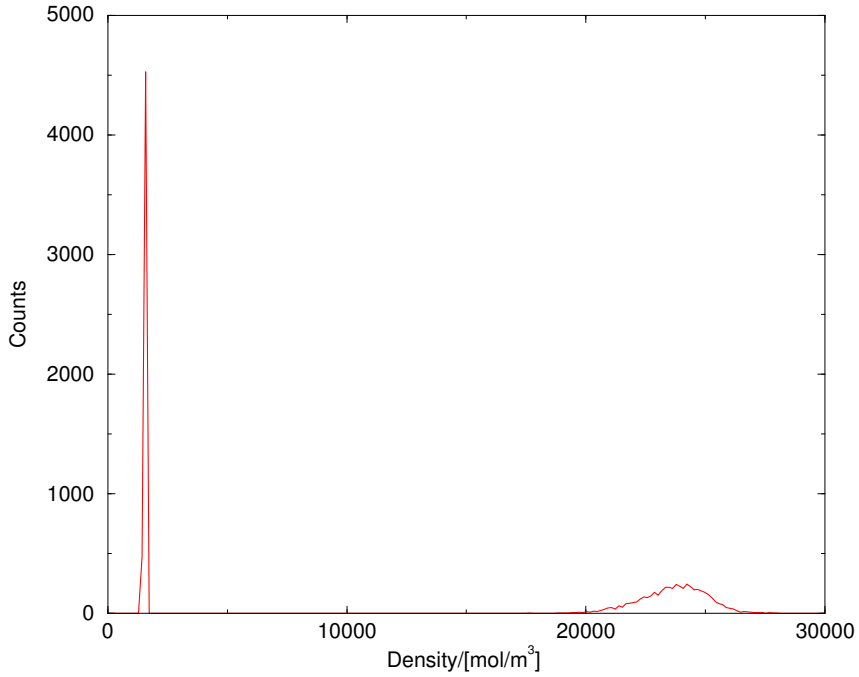


Figure 7.12.: BBV 250 K: Density Histogram.

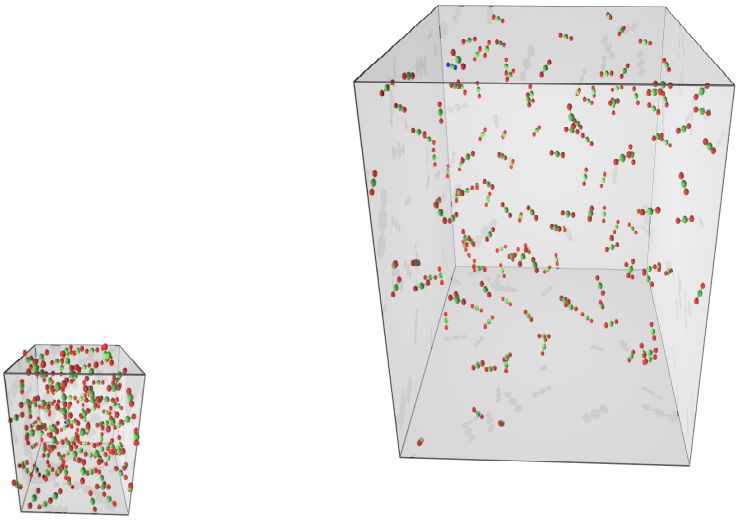


Figure 7.13.: BBV 250 K Snapshot.

7.3.5. 260 K

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
1849	22598	29.8
1852	22610	30.1
1885	22737	30.3
2015	22529	31.0
2214	22697	33.4
2099	22636	32.0
2008	22363	31.3

Table 7.14.: BBV 260 K: Results.

Property	**Value**
ρ_g:	1989 ± 52 $\frac{mol}{m^3}$
ρ_l:	22596 ± 46 $\frac{mol}{m^3}$
P:	31.1 ± 0.5 bar
$\rho_{g,\text{exp}}$:	1465.1 $\frac{mol}{m^3}$
$\rho_{l,\text{exp}}$:	22709.7 $\frac{mol}{m^3}$
P_{exp}:	24.194 bar

Table 7.15.: BBV 260 K: Averages.

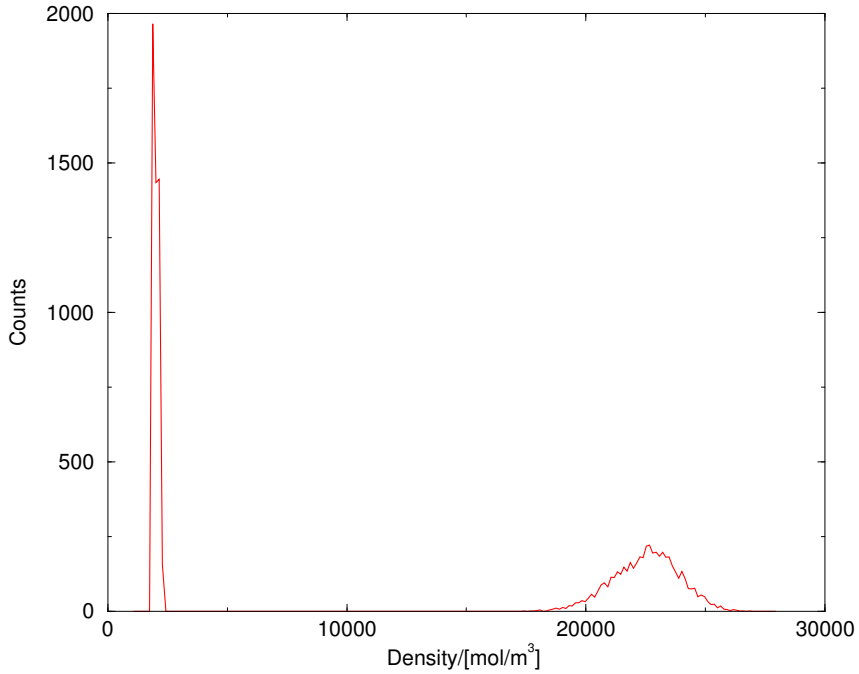


Figure 7.14.: BBV 260 K: Density Histogram.

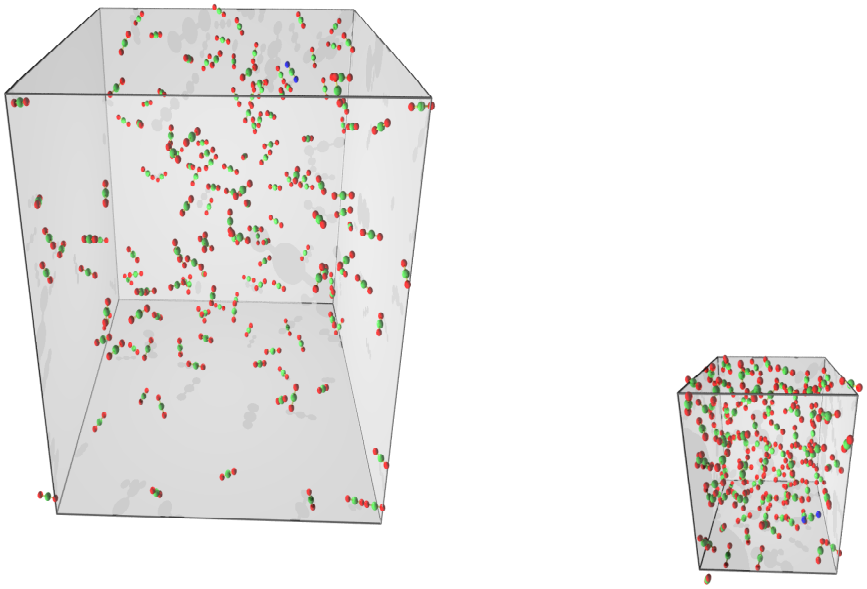


Figure 7.15.: BBV 260 K Snapshot.

7.3.6. 270 K

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
3007	21489	42.9
2978	21336	43.8
3017	21468	43.6
2992	21212	43.6

Table 7.16.: BBV 270 K: Results.

Property	**Value**	
ρ_g:	2999 ± 9	$\frac{mol}{m^3}$
ρ_l:	21376 ± 64	$\frac{mol}{m^3}$
P:	43.5 ± 0.2	bar
$\rho_{g,\mathrm{exp}}$:	2008.0	$\frac{mol}{m^3}$
$\rho_{l,\mathrm{exp}}$:	21504	$\frac{mol}{m^3}$
P_{exp}:	32.034	bar

Table 7.17.: BBV 270 K: Averages.

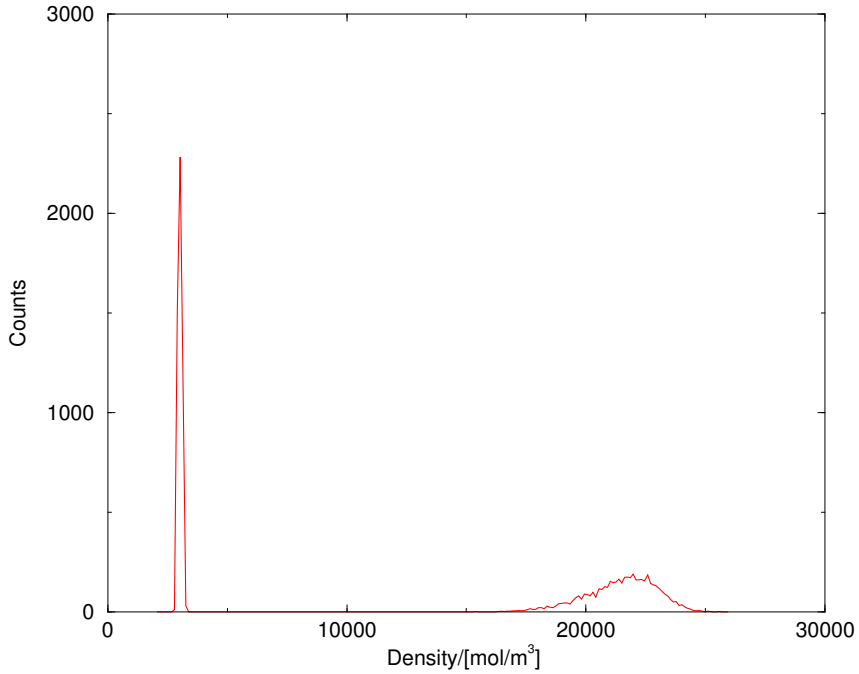


Figure 7.16.: BBV 270 K: Density Histogram.

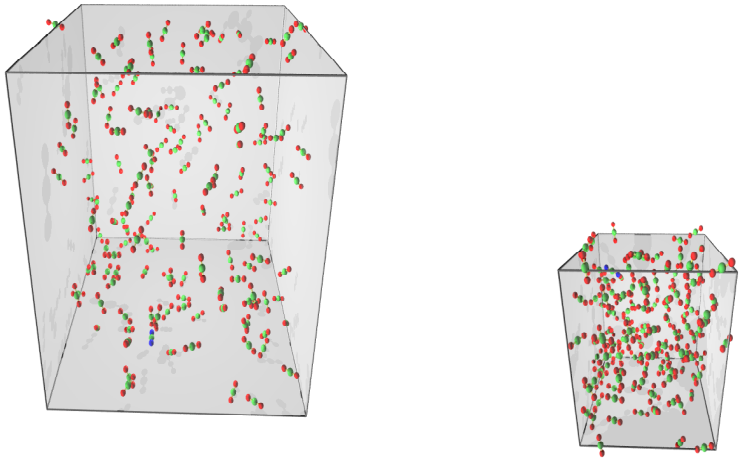


Figure 7.17.: BBV 270 K Snapshot.

7.3.7. 280 K

The temperature is already close to the critical temperature and the free energy difference between both phases is getting smaller. Phase changes in the two subsystems occur. This can be seen in the density histogram 7.18. The fluctuations get larger and the accuracy of the results gets lower. The average densities of both phases had to be extracted from the histograms.

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
3940	20000	54.2
4000	20100	54.2
3700	19800	54.9
3600	20100	56.6

Table 7.18.: BBV 280 K: Results.

Property	**Value**
ρ_g:	3810 ± 95 $\frac{mol}{m^3}$
ρ_l:	20000 ± 71 $\frac{mol}{m^3}$
P:	55.0 ± 0.6 bar
$\rho_{g,\text{exp}}$:	2763.7 $\frac{mol}{m^3}$
$\rho_{l,\text{exp}}$:	20091 $\frac{mol}{m^3}$
P_{exp}:	41.595 bar

Table 7.19.: BBV 280 K: Averages.

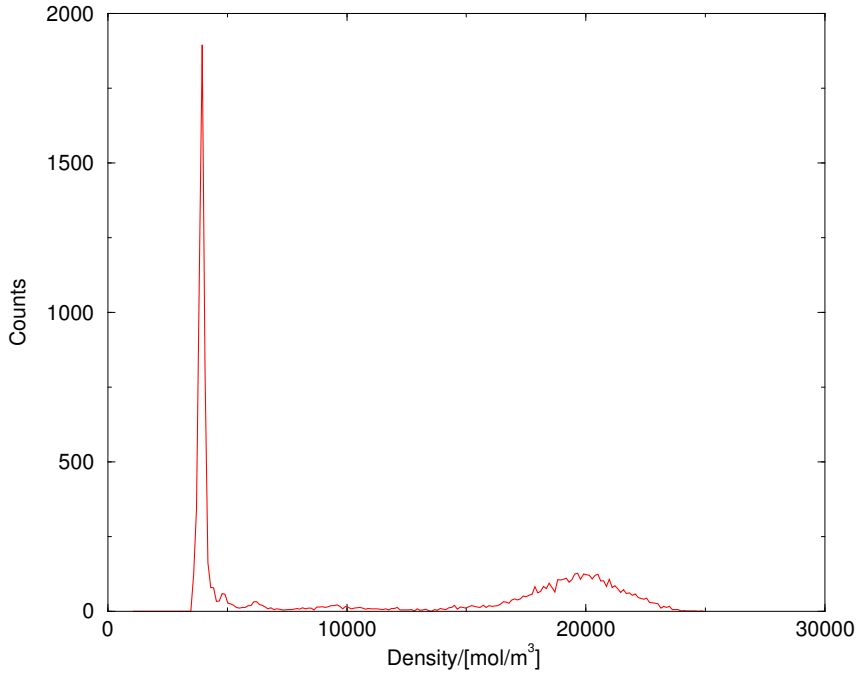


Figure 7.18.: BBV 280 K: Density Histogram.

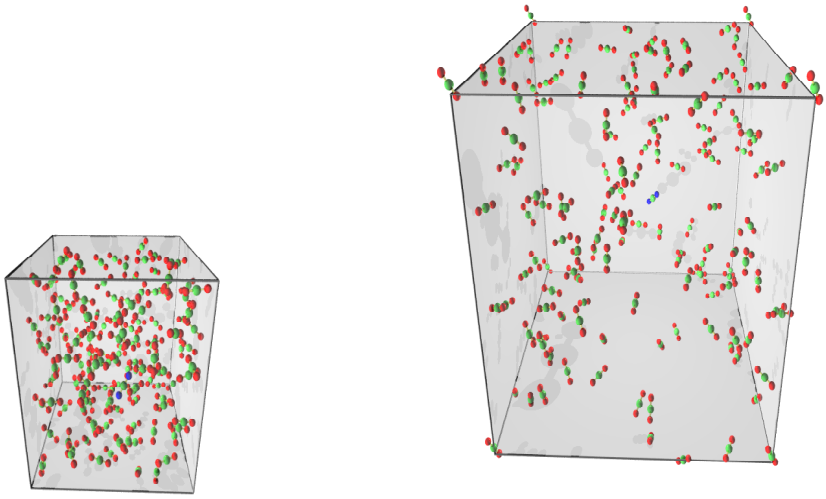


Figure 7.19.: BBV 280 K Snapshot.

7.3.8. 290 K

Getting closer to the critical point, the density histogram shows that more phase changes occur, the points between the two phases have a higher weight.

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
5100	18500	71.8
5500	18800	71.7
5600	18000	72.3
5400	18000	70.6
5400	18500	73.7

Table 7.20.: BBV 290 K: Results.

Property	**Value**	
ρ_g:	5400 $\pm$ 84	$\frac{mol}{m^3}$
ρ_l:	18360 $\pm$ 157	$\frac{mol}{m^3}$
P:	72.0 $\pm$ 0.5	bar
$\rho_{g,\text{exp}}$:	3902.4	$\frac{mol}{m^3}$
$\rho_{l,\text{exp}}$:	18301.3	$\frac{mol}{m^3}$
P_{exp}:	53.152	bar

Table 7.21.: BBV 290 K: Averages.

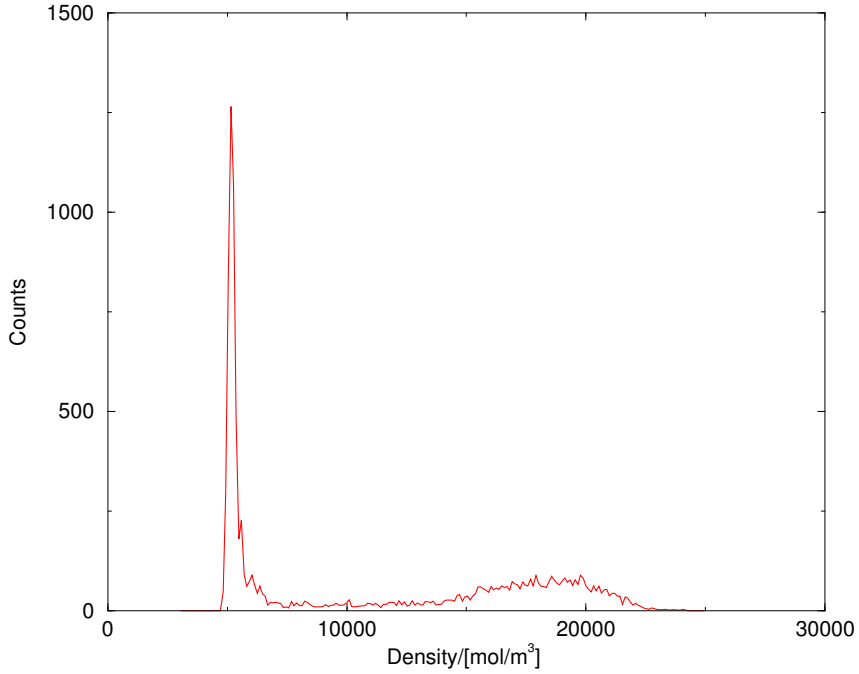


Figure 7.20.: BBV 290 K: Density Histogram.

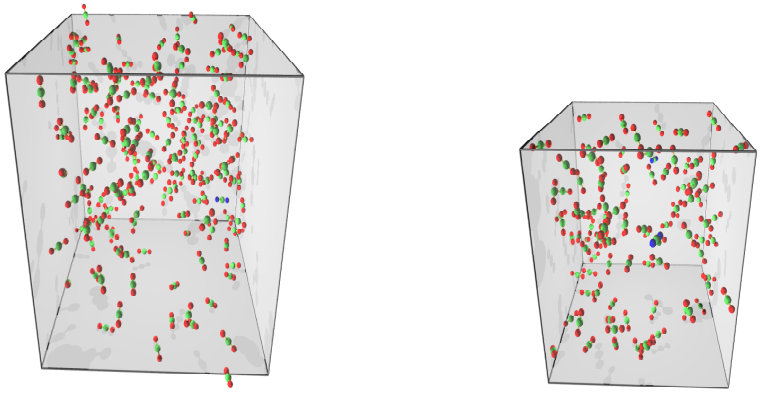


Figure 7.21.: BBV 290 K Snapshot.

7.3.9. 300 K

The 300 K simulations are in the vicinity of the critical point. Often no clear phase separation was seen in the density histograms, only the vapor maximum with a wide liquid distribution was observed. To improve the statistics, several simulations were done at this temperature. The selected runs with a clear phase separation are listed below. The standard deviation is high.

$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
7200	15000	86.8
7700	13500	89.6
8000	15000	87.3
8600	14000	88.1
8500	16000	90.8
9000	15000	89.1

Table 7.22.: BBV 300 K: Results.

Property	**Value**
ρ_g:	8167 $\pm$ 269 $\frac{mol}{m^3}$
ρ_l:	14750 $\pm$ 359 $\frac{mol}{m^3}$
P:	88.6 $\pm$ 0.6 bar
$\rho_{g,\mathrm{exp}}$:	6097.6 $\frac{mol}{m^3}$
$\rho_{l,\mathrm{exp}}$:	15458 $\frac{mol}{m^3}$
P_{exp}:	67.095 bar

Table 7.23.: BBV 300 K: Averages.

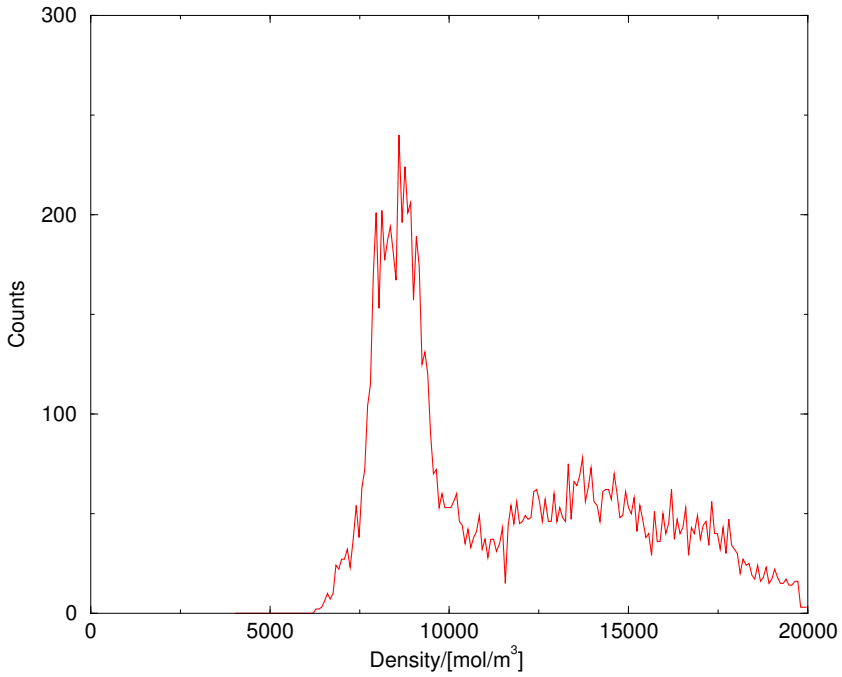


Figure 7.22.: BBV 300 K: Density Histogram.

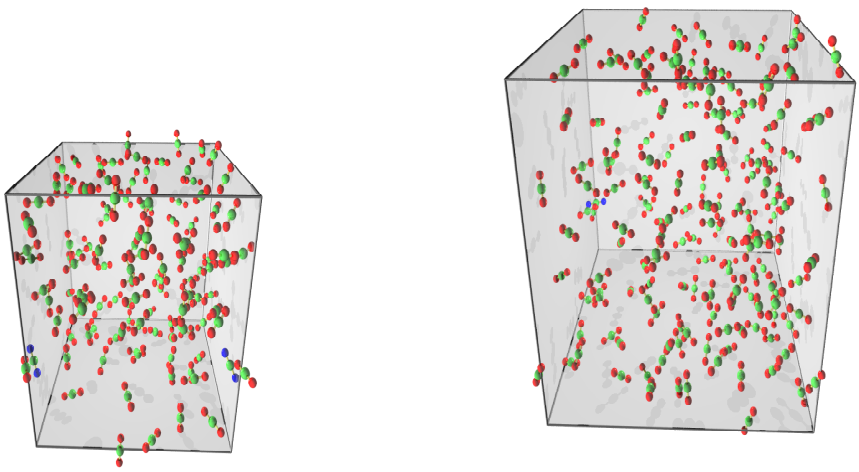


Figure 7.23.: BBV 300 K Snapshot.

7.3.10. 310 K

The 310 K simulation is above the critical temperature in the supercritical region. The density histogram (figure 7.24) shows a single phase which is much broader than the liquid phase at 280 K shown in figure 7.3 on page 78.

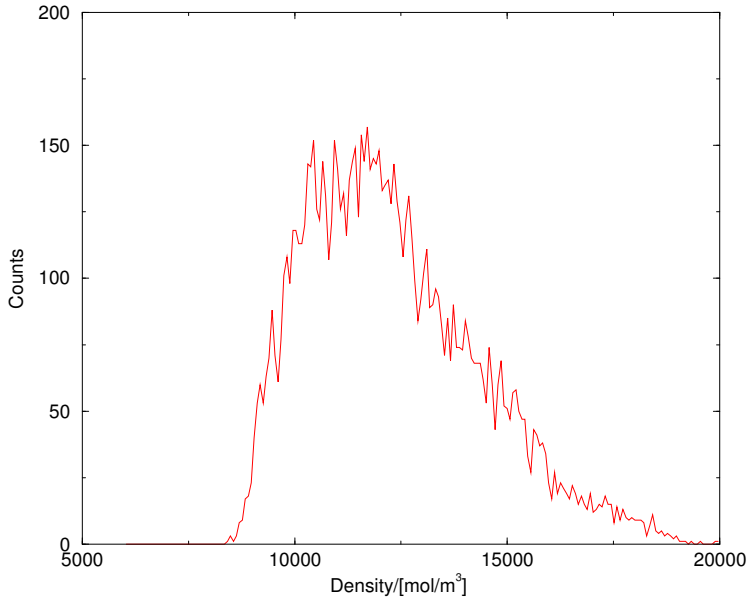


Figure 7.24.: BBV 310 K: Density Histogram.

7.3.11. Summary

The liquid-vapor equilibria were calculated from 220 K up to the critical point. Table 7.24 lists all equilibrium values and figure 7.25 shows the density histograms at different temperatures.

Temperature/$[K]$	$\rho_g/[\frac{mol}{m^3}]$	$\rho_l/[\frac{mol}{m^3}]$	$P/[bar]$
220	421 ± 6	26851 ± 38	6.9 ± 0.1
230	592 ± 4	25924 ± 33	9.8 ± 0.1
240	850 ± 11	24979 ± 12	14.3 ± 0.2
250	1502 ± 21	23758 ± 68	23.4 ± 0.3
260	1989 ± 52	22596 ± 46	31.1 ± 0.5
270	2999 ± 9	21376 ± 64	43.5 ± 0.2
280	3810 ± 95	20000 ± 71	55.0 ± 0.6
290	5400 ± 84	18360 ± 157	72.0 ± 0.5
300	8167 ± 269	14750 ± 359	88.6 ± 0.6

Table 7.24.: BBV Gibbs Ensemble Molecular Dynamics: Liquid-Vapor Results.

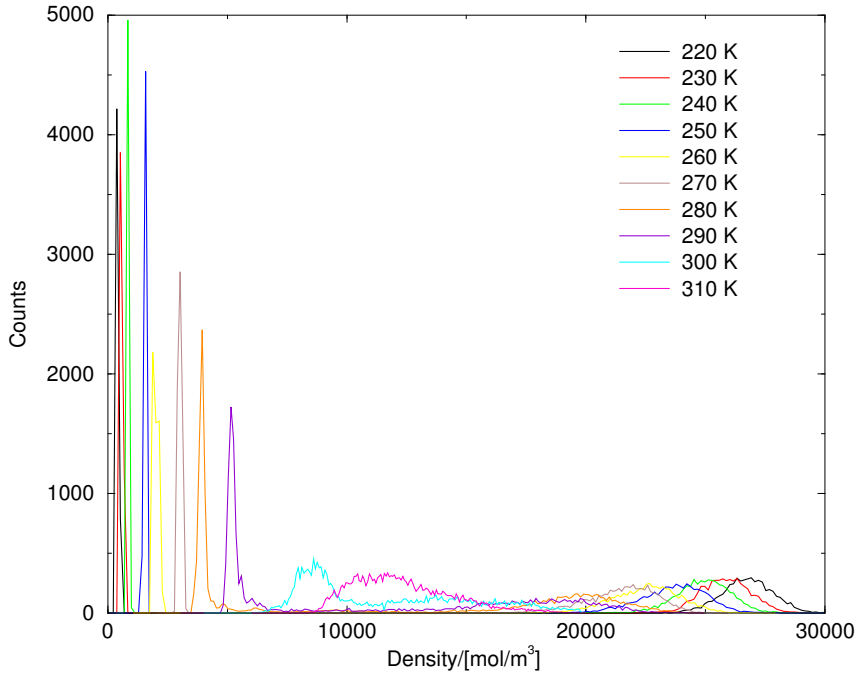


Figure 7.25.: BBV Gibbs Ensemble Molecular Dynamics: Liquid-Vapor Normalized Density Histograms.

7.4. Fitting the Critical Properties

The simplest method to get the critical temperature directly is to increase the temperature until a phase separation can no longer be observed. The problem with this method is that the correlation length diverges at the critical point. Therefore a finite size system used in molecular dynamics is not able to sample a macroscopic effect. Upon getting closer to the critical point the correlation length increases and at a certain temperature the correlation length is equal to the box length. After this crossover temperature the obtained values are no longer accurate enough. Even experiments are limited by the size of the vessel. External disturbances, *e.g.* the gravity, have to be zero to get accurate results.

The Gibbs molecular dynamics simulations deliver the density of the liquid ρ_l and the gas ρ_g at a certain temperature. To obtain the critical temperature T_c and density ρ_c, two laws are used[24, p. 216][62, 71]. Panagiotopoulos showed in [62] that system size effects on the coexistence curve are small and that the crossover temperature is close to the critical temperature. Therefore both laws can be used, even for small systems, to extrapolate the missing near critical values.

The first one is called the law of rectilinear diameters.

$$\frac{\rho_l + \rho_g}{2} = \rho_c + A(T - T_c) \tag{7.8}$$

The second one is the scaling law.

$$\rho_l - \rho_g = B(T - T_c)^{\beta} \tag{7.9}$$

All variables are colored in blue and all fit parameters in red. A and B are constants without physical meaning obtained from the fit. β is called the critical exponent (so-called Ising parameter), which is usually 0.32 for a three dimensional system[24]. Sengers and coworkers[71] gave a value of approximately 0.35 for CO_2 and N_2O.

The experience with the Monte Carlo method points out that an accuracy of $\pm 1\%$ and better is possible for the critical temperature[62]. The accuracy of the critical density is lower.

The experimental critical values are shown in the next table:

Property	**Value**
Temperature:	304.21 K
Pressure:	73.825 bar
Density:	10590 $\frac{mol}{m^3}$

Table 7.25.: CO_2 Experimental Critical Values[6].

7.4.1. Results

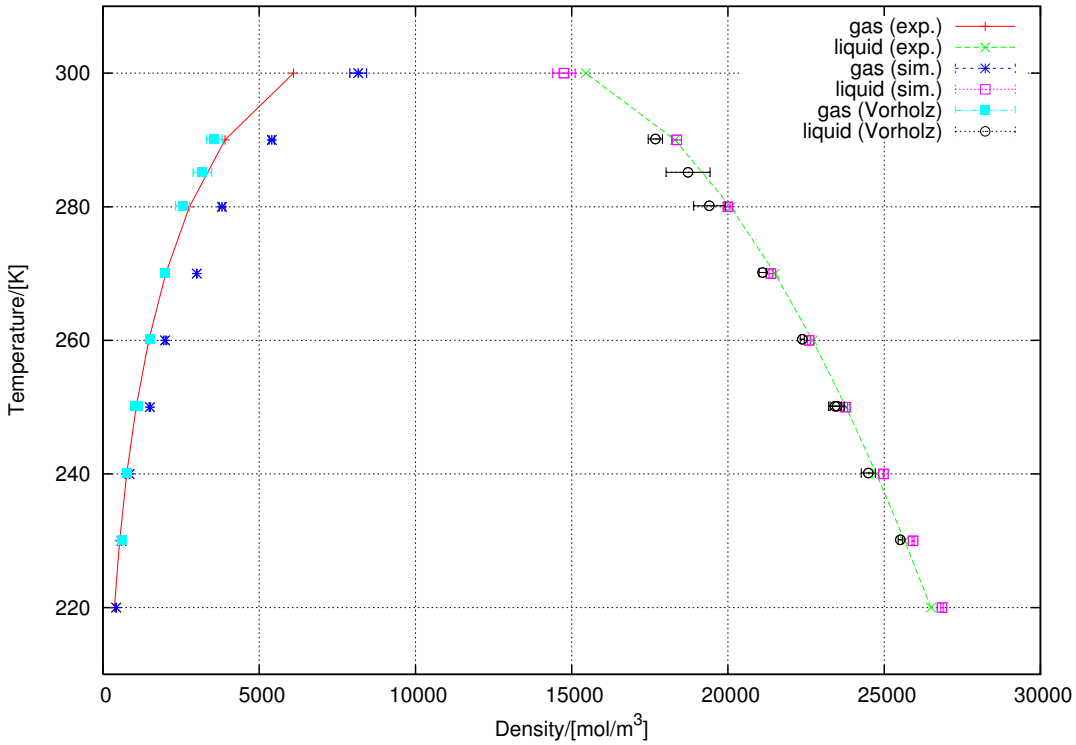


Figure 7.26.: BBV Liquid-Vapor Coexistence Curve.

The values from Vorholz *et al.* were obtained by Gibbs ensemble Monte Carlo calculations using the EPM2 CO_2 potential[93]. This potential was optimized to reproduce the critical values. The calculated critical values are[31] 313.4 ± 0.7 K, 10309 ± 98 mol m^{-3} and 76.5 ± 4.5 bar.

At lower temperatures, all calculated values are in good agreement with experiment. The differences increase until the critical point is reached. The liquid BBV values reproduce the experimental values very well. The vapor density is always larger than the experimental values. Several of the high temperature points were recalculated with a larger system of 432 molecules but the same average values were obtained. No improvement was seen with the larger cut-off radius. The higher critical density is in agreement with the minimum in the rotational relaxation time (see figure 8.1 on page 107). Simulations using two equal densities were done in the density region between and below the experimental and the simulation vapor values. No phase separation was seen and stable gas phases were sampled. Therefore the deviation from the experiment is not an error in the Gibbs ensemble calculations, it is a characteristic of the pair potential. The fits below show that the scaling laws are valid for the simulation values.

The BBV pressures are always above the experimental ones. The values are between 10 % and 40 % higher because of the higher average densities. Single phase values were compared in section 7.1.1 on page 75.

Exchange Duration

Temperature/$[K]$	**Duration**/$[fs]$	**Failed Duration**/$[fs]$	**Duration**/$\rho_l/\left[\frac{fs\,m^3}{mol}\right]$
220	1580	250	0.059
230	1440	249	0.056
240	1336	248	0.053
250	1332	235	0.056
260	1260	233	0.056
270	1205	210	0.056
280	1112	198	0.055
290	1033	187	0.056
300	1044	171	0.071

Table 7.26.: CO_2 BBV Gibbs Ensemble Molecular Dynamics Exchange Duration.

The duration of particle exchanges scales linearly with the liquid density. The exchange algorithm is very efficient.

7.4.2. Scaling Law Fit

The scaling law was fitted with the function

$$T = C\left(\frac{\rho_l - \rho_g}{\rho_0}\right)^{\frac{1}{\beta}} + T_c \tag{7.10}$$

including β as a fit parameter. The errors in the densities were transformed to errors in the temperature. ρ_0 was added to the fit function, where ρ_0 is 1 $\mathrm{mol\,m^{-3}}$, to give the parameter C a well defined dimension. The best fit for the simulation data was obtained with a β value close the experimental one of 0.35[71]. The experimental points with a β close to the Ising value of 0.32 give a critical temperature which is in agreement with the literature value in table 7.25.

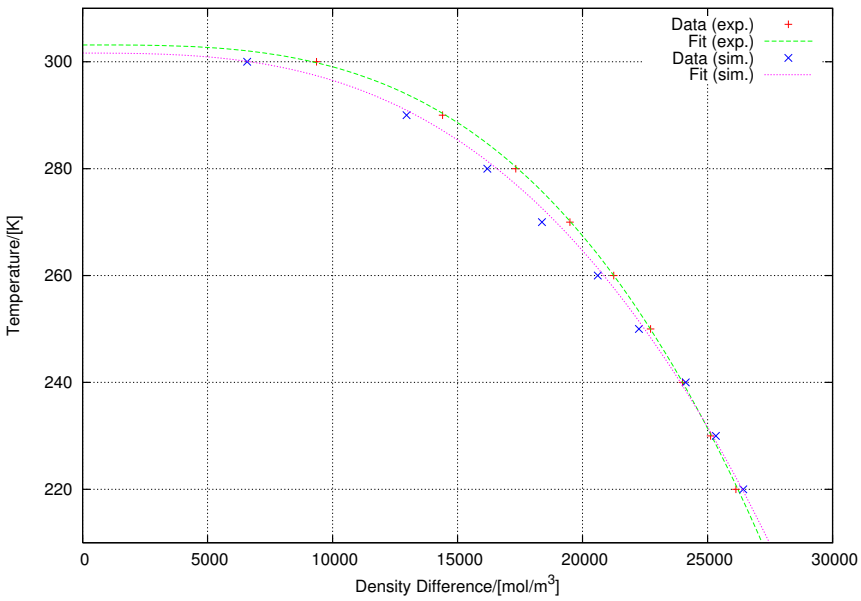


Figure 7.27.: BBV Scaling Law Fit.

Property	**Value**	
$T_{c,\mathrm{exp}}$:	303.1 ± 1.3	K
β_{exp}	0.32 ± 0.01	
C_{exp}:	-1.30x10^{-12} ± 1.62x10^{-12}	K
$T_{c,\mathrm{sim}}$:	301.6 ± 0.3	K
β_{sim}	0.350 ± 0.004	
C_{sim}:	-1.9x10^{-11} ± 6.3x10^{-12}	K

7.4.3. Rectilinear Diameter Fit

The critical temperature, known from the scaling law, was used for the rectilinear diameter fit. The errors of the critical temperatures are small and were not accommodated.

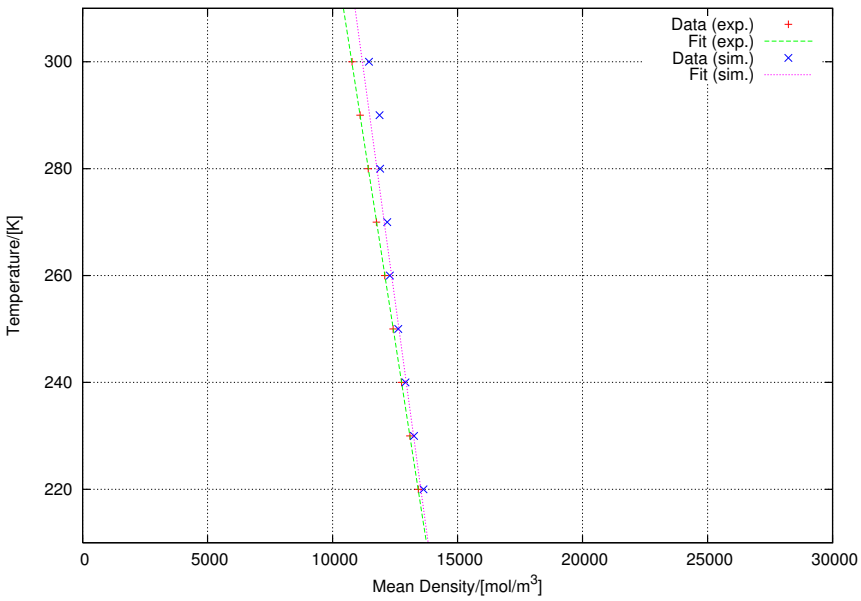


Figure 7.28.: BBV Rectilinear Diameter Fit.

Property	Value	
$\rho_{c,\mathrm{exp}}$:	10664 $\pm$ 0.6	$\frac{mol}{m^3}$
A_{exp}:	-33.2 $\pm$ 0.01	$\frac{mol}{K\,m^3}$
$\rho_{c,\mathrm{sim}}$:	11145 $\pm$ 34	$\frac{mol}{m^3}$
A_{sim}:	-29.2 $\pm$ 0.5	$\frac{mol}{K\,m^3}$

7.4.4. Pressure Fit

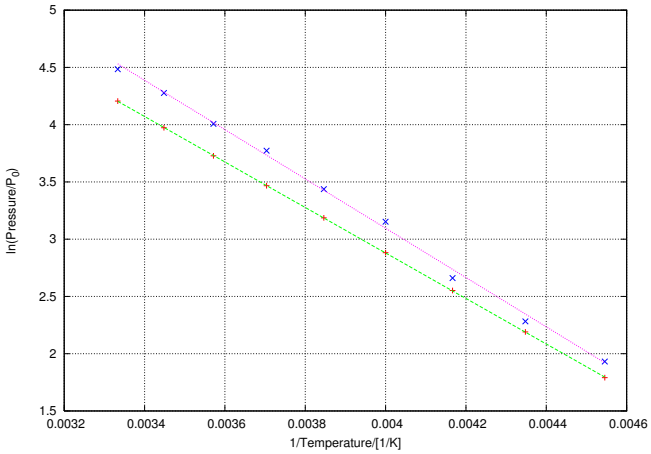


Figure 7.29.: BBV Vapor Pressure Fit.

The vapor pressure was fitted to the Clausius-Clapeyron equation[7],

$$\ln \frac{P}{P_0} = -\frac{\Delta H_{\text{vap}}}{RT} + D \tag{7.11}$$

giving the following values for the pressure at the critical temperature and the vaporization enthalpy:

Property	Value
$P_{c,\text{exp}}$:	71.6 ± 2.2 bar
$\Delta H_{\text{vap,exp}}$	16.52 ± 0.02 $\frac{kJ}{mol}$
$P_{c,\text{sim}}$:	96.5 ± 4.2 bar
$\Delta H_{\text{vap,sim}}$	17.90 ± 0.07 $\frac{kJ}{mol}$

All errors in the fit parameters and in the temperature were used to calculate the critical pressure. P_0 was set to 1 bar. The fitted experimental value is in the range of the literature value (see table 7.25). The enthalpy of sublimation is 26.25 ± 0.09 kJ mol^{-1}[6].

The inverse temperature plotted against the logarithm of the pressure is not an exact linear function of the simulated values.

7.4.5. Critical Properties

Property	Value
$T_{c,\text{sim}}$:	301.6 ± 0.3 K
$T_{c,\text{exp}}$:	303.1 ± 1.3 K
$\rho_{c,\text{sim}}$:	11145 ± 34 $\frac{mol}{m^3}$
$\rho_{c,\text{exp}}$:	10664 ± 0.6 $\frac{mol}{m^3}$
$P_{c,\text{sim}}$:	96.5 ± 4.2 bar
$P_{c,\text{exp}}$:	71.6 ± 2.2 bar

Table 7.27.: BBV and Experimental Critical Properties (Fitted Values)

8. Critical Properties and Fluid Structure

Supercritical fluids are used as solvents for extraction processes and chemical reactions. The temperature must be above the critical temperature to be in the supercritical region. Often the critical region is defined to start where the pressure is also larger than the critical pressure. Here, the first definition is used where the temperature has to be above the critical temperature. The advantage of the critical region is that physical properties can be optimized by varying the pressure and temperature over a wide range without any phase changes. No liquid phase exists anymore above the critical point. Close to the critical point, the correlation length, the width of homogeneous fluid behavior, gets larger and the correlation length is infinite at the critical point. The critical opalescence is a consequence of the divergence. Due to regions of different densities, light of the same wavelength as the diameter of a density region, is strongly scattered. At the critical point CO_2 is perfectly reflecting and in the vicinity it has a pale red color[65]. The speed of sound drops to a minimum. Supercritical CO_2 has a similar solvent power as petrol. It has the ability to dissolve fluorinated organic compounds. Supercritical water is very reactive, it can be used to destroy nerve-gas weapons.

Mathematically, the critical point is defined by the following two equations.

$$\left(\frac{\partial P}{\partial \rho}\right)_T = 0 \tag{8.1}$$

$$\left(\frac{\partial^2 P}{\partial \rho^2}\right)_T = 0 \tag{8.2}$$

The first and second derivatives are zero, higher derivatives are non-zero. In other words, the critical isotherm has an inflection point and the isothermal compressibility

$$\left(\frac{\partial \rho}{\partial P}\right)_T = \infty \tag{8.3}$$

becomes infinite. The pressure in the vicinity of the critical point is independent of the density.

If the ideal gas equation of state is inserted into both equations, no reasonable solution is possible: the ideal gas has no critical point. Using the van der Waals equation of state (equation

(8.4)), the critical point can be parameterized but properties of the van der Waals equation are significantly different to experimental values. The two additional parameters of the van der Waals equation are not sufficient to fully describe the behavior of the system in this region.

$$P = \frac{RT}{V_m - b} - \frac{a}{V_m^2} \tag{8.4}$$

$$V_{m,c} = 3\,b \tag{8.5}$$

$$T_c = \frac{8\,a}{27\,b\,R} \tag{8.6}$$

$$P_c = \frac{a}{27\,b^2} \tag{8.7}$$

In the Gibbs ensemble chapter 7, the two-phase behavior of CO_2 up to the critical point was shown. Above the critical temperature, the beginning of the critical region, the fluid structure is not as homogeneous as in a pure gas or liquid phase. The critical region can be divided into a more gas-like and a more liquid-like part. Some degrees K above the critical temperature the density fluctuations are still appreciable and regions of different densities exist.

In this chapter, it will be shown how the divergence of the correlation length was used to find the critical point, and how average structural properties and the dynamic fluid behavior manifest themself in the critical region. All simulations were done in the microcanonical ensemble with at least 512 molecules. To get more accurate values, one should use a thermostat. The results shown here were calculated before the Gibbs ensemble method was used to fit the critical properties. Therefore the critical values were estimated by near critical isotherms. Finite size effects play an important role in this region. The estimated values for the flexible BBV potential were:

Property	**Value**	
Temperature:	300 ± 5	K
Pressure:	84.7 ± 6	bar
Density:	11500 ± 500	$\frac{mol}{m^3}$

The errors are large, but the simple single system approach was able to deliver the correct values. These values were obtained first from a 304 K isotherm using self diffusion, reorientational, rotational, thermal conductivity and viscosity time auto-correlation functions. A divergence in the isotherms was seen in the viscosity, thermal conductivity[76] and the rotational relaxation time[55]. The rotational relaxation time, which is calculated by second order Legendre polynomials, clearly showed a minimum at the critical density. The rotational relaxation of supercritical CO_2 1 % above the critical temperature was studied by Adams and Siavosh-Haghighi[1] using the scaled EPM2 potential[31]. A flat minimum in the vicinity of the critical region, with values in the same range as shown in the next figure, was obtained. Raman spectra exhibit the same behavior[59].

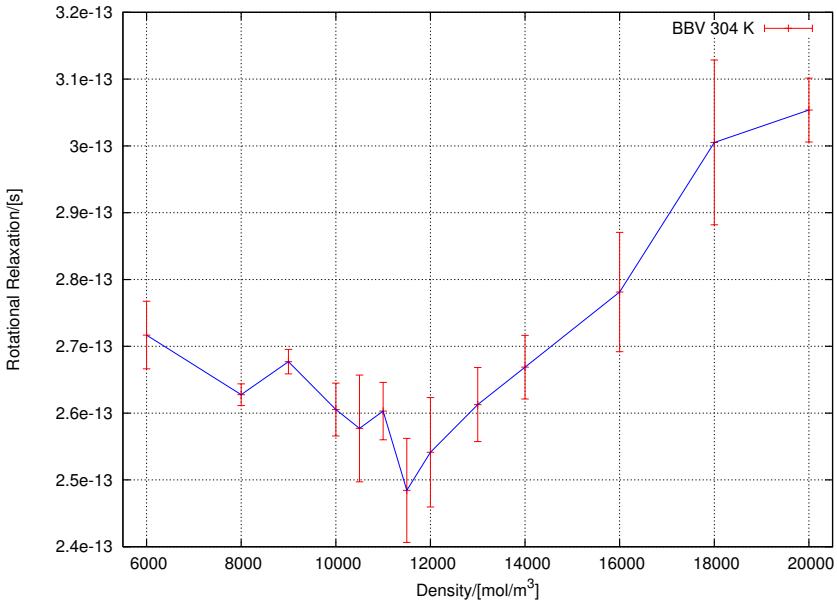


Figure 8.1.: CO_2 BBV 304 K Isotherm: Rotational Relaxation Time.

By varying the temperature at the critical density, the critical temperature was obtained.

The methods discussed here were developed to describe the behavior of supercritical fluids. Tucker wrote in a review[84] about the effects caused by inhomogeneities of supercritical fluids. Most methods shown in this chapter are based on this review and succeeding publications. The main difference in this work is that an *ab initio* potential is used in a 3-D space (instead of argon, Lennard-Jones and 2-D). Both extensions limit the maximum accessible size of the system. In reference [48] Tucker *et al.* used 1152 atoms in a quadratic system. If the same box length would be used in the 3-D system, then the number of particles would be 39304. One should be careful in comparing fluctuations of 2-D and 3-D systems[26]. They did the simulations in the NVE ensemble, with velocity rescalings every 33rd time step. Therefore no canonical temperature fluctuations were sampled. The velocity rescaling could introduce a significant error in the density distribution.

8.1. Local Density Enhancement

Large density fluctuations in the vicinity of the critical point and a diverging correlation length at the critical point are the main reasons for the observed phenomena in the critical region. To study the density fluctuations and average density enhancements, an atom-based density given by a sphere is introduced. The local density enhancement is a static property of supercritical fluids.

The average number of molecules in a sphere of radius r_{ck} is a fixed number for an ideal gas:

$$n_0 = \frac{N}{V}\frac{4}{3}\pi r_{ck}^3 \tag{8.8}$$

In a real liquid, the spatial distribution is not homogeneous. The radial pair distribution function $g(r)$, the g-function, shows the average probability of a pair distance relative to the ideal gas. The ideal gas always has a probability of one.

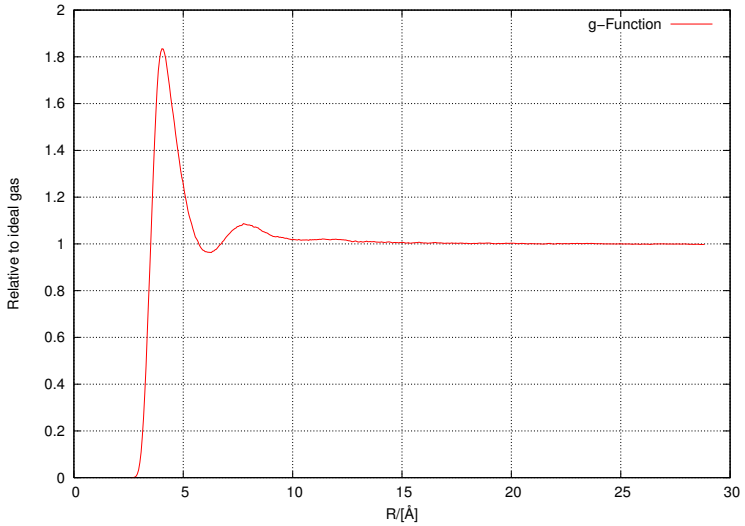


Figure 8.2.: CO_2 BBV Carbon-Carbon g-Function at 303 K and 11500 mol m^{-3}.

For molecules, several g-functions between different atoms of the dimers exist. Figure 8.2 shows the carbon-carbon g-function which is equal to the center of mass pair distribution function. Mean values are known from neutron and X-ray scattering experiments[16].

8.1.1. Graphical Analysis

To show that zones of different local densities exist, three different categories are defined: low, average and high density regions. The number of molecules n_{ck} in a sphere of radius r_{ck} around the center of mass of each molecule are calculated. In the case of a single component system the center molecule is also counted.

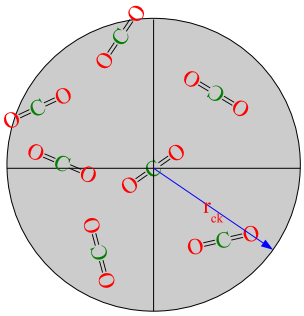


Figure 8.3.: Local Density Cut-Off Sphere.

Each molecule is in one of the three regions:

$$\begin{array}{lcccll} n_{ck} & > & n_0 + \delta & & & \textbf{high} \\ n_0 - \delta & \leq & n_{ck} & \leq & n_0 + \delta & \textbf{average} \\ n_{ck} & < & n_0 - \delta & & & \textbf{low} \end{array}$$

The same threshold ($\pm$ 15 %) as in the original paper[48] was used:

$$\delta = 0.15\, n_0 \tag{8.9}$$

Different snapshots of the following simulation were used to visualize the three density regions (see appendix B.3.4 on page 153):

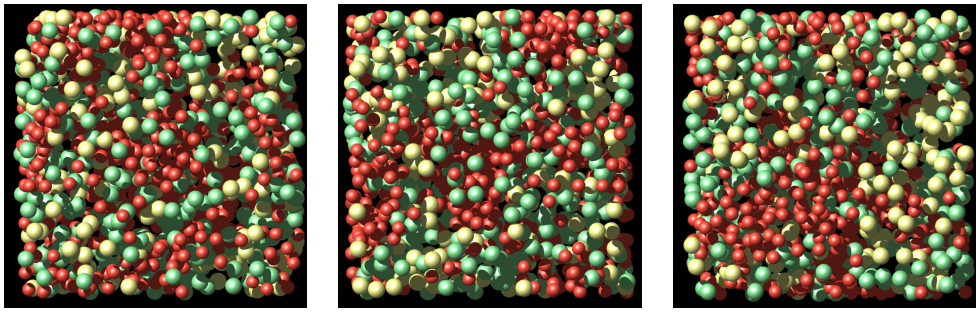
Property	**Value**
Temperature:	302.8 K
Density:	11500 $\frac{mol}{m^3}$
Box length:	57.7 Å
Molecules:	1331

To simplify the graphs, each molecule was reduced to a sphere. In the following graphs, the regions at three different radii are shown. The simulation box is presented from all three directions. On the pictures with all three regions on page 111, the high region is at the top, the average region in the middle and the low region at the bottom.

The graphs show that different density regions exist. Plots of several snapshots of the same run showed that the number of molecules in the three regions is nearly constant with time. The graphs indicate that the high density region is dominant at small r_{ck} values. This phenomenon

is called local density enhancement. At larger r_{ck} values the bulk density is reached. If experiments are done, most properties are related to pair distributions. The surrounding molecules interact with the central molecule, changing the interactions and therefore the characteristics of the fluid. The simulations shown here were done using a single component system, mixtures show solvent effects given by different solubilities in the different regions.

6 Å



front **left** **top**

$$r_{ck} = 6\,\text{Å}$$

$$\begin{aligned} n_{lo} &= 231 \\ n_{av} &= 399 \\ n_{hi} &= \underline{701} \end{aligned}$$

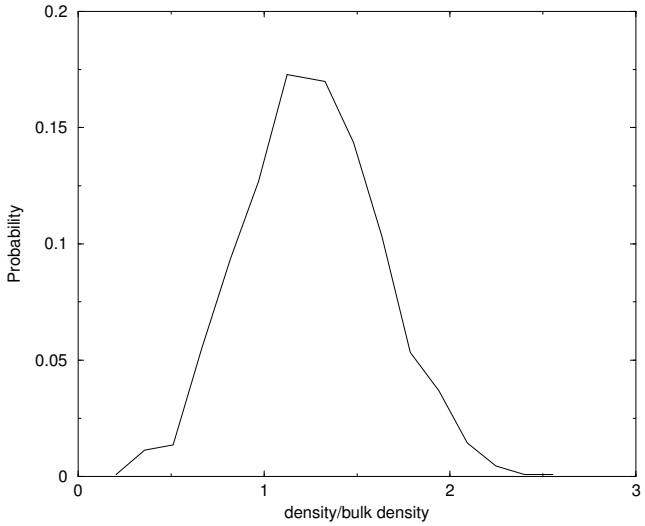
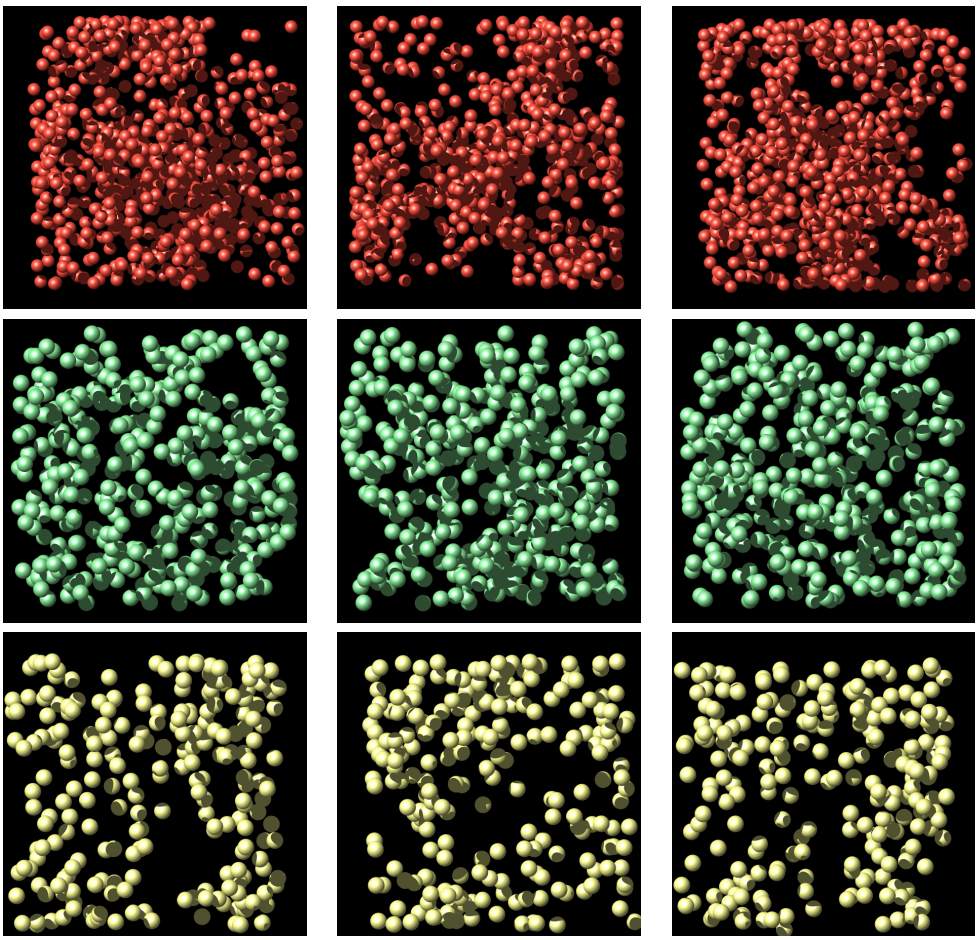


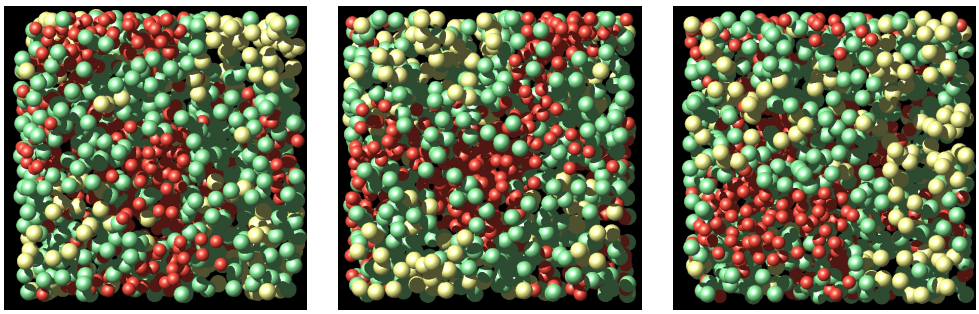
Figure 8.4.: Relative Density Probability $r_{ck} = 6\,\text{Å}$.



front **left** **top**

At 6 Å most molecules belong to the high density region.

10 Å



front **left** **top**

$$r_{ck} = 10\,\text{Å}$$

$$\begin{aligned} n_{lo} &= 213 \\ n_{av} &= \underline{607} \\ n_{hi} &= 511 \end{aligned}$$

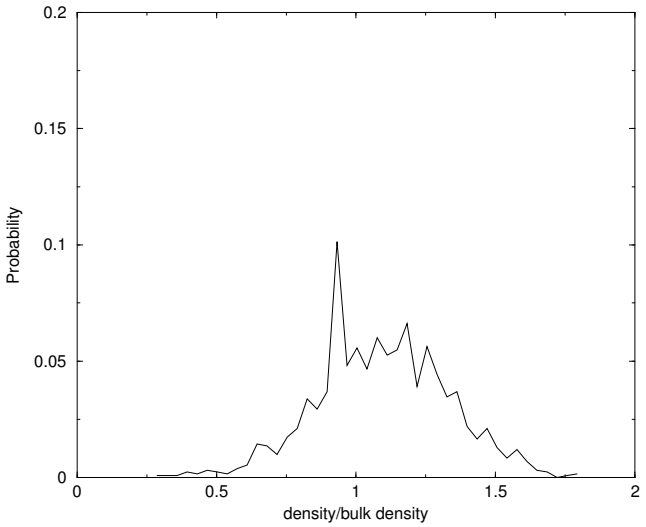
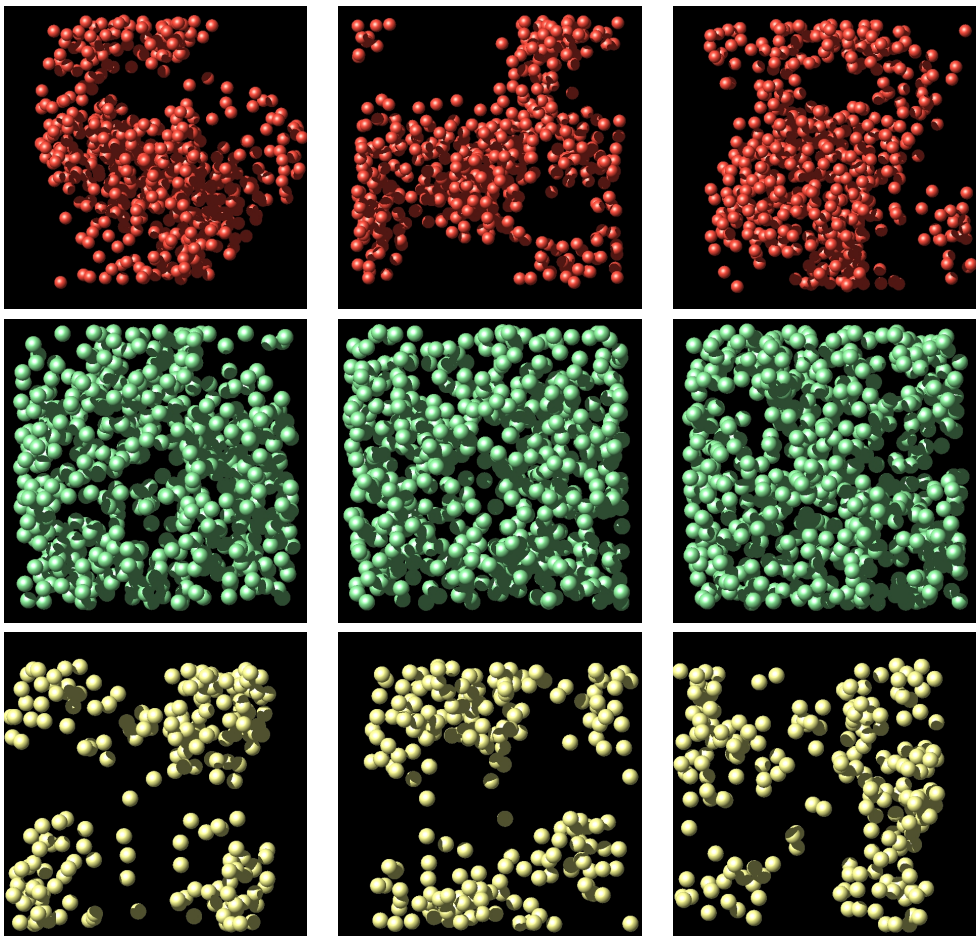


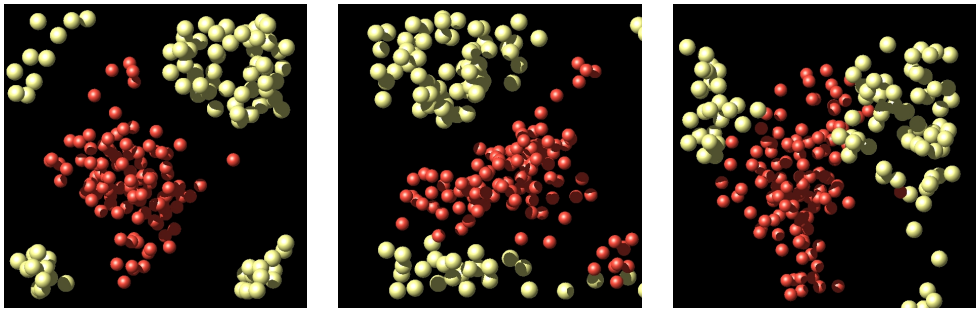
Figure 8.5.: Relative Density Probability $r_{ck} = 10\,\text{Å}$.



front **left** **top**

At 10 Å most molecules are in the average density region. The high density region contains more than twice the number of low density molecules.

20 Å



front **left** **top**

Most of the molecules are in the average density region (not shown). It is clearly visible that

the high density region is in the center of the box. The low density region is at the corner.

$$
\begin{aligned}
r_{ck} &= 20\,\text{Å} \\
\\
n_{lo} &= 96 \\
n_{av} &= \underline{1122} \\
n_{hi} &= 113
\end{aligned}
$$

The number of low and high density molecules are equal.

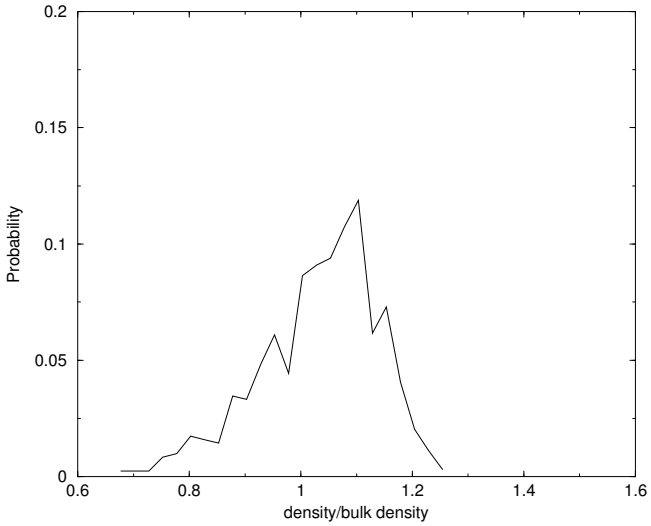
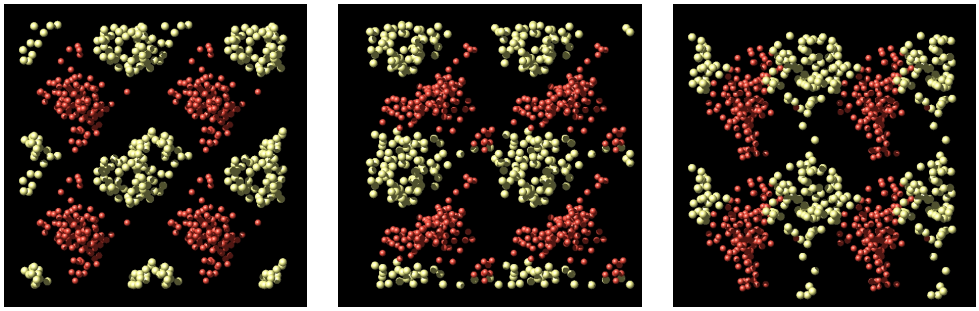


Figure 8.6.: Relative Density Probability $r_{ck} = 20\,\text{Å}$.



front **left** **top**

In the last three images, two simulation boxes are connected in each direction, to construct a cubic box of 8 unit cells. The regions at the corner of the unit cell build now a closed region. Both the low and high density regions form droplets with molecules of average density in between.

Total and Partial g-Functions

The three regions are further analyzed with local pair distribution functions. Two different types of g-functions are used: total pair distribution functions show the fluid structure around each of the three density regions, partial g-functions show the distribution of one of the three density regions around a density species[27]. Six different types of partial g-functions are possible

low - low	low - average	low - high
average - average	average - high	high - high

where low-average is equal to average-low. The low-high partial g-function shows the distribution of all high region molecules around the low region molecules, where the low total g-function shows the distribution of all molecules around the low region molecules.

The following graphs show both types of g-functions. The g-functions were calculated from the data of 10 snapshots from the same run.

6 Å

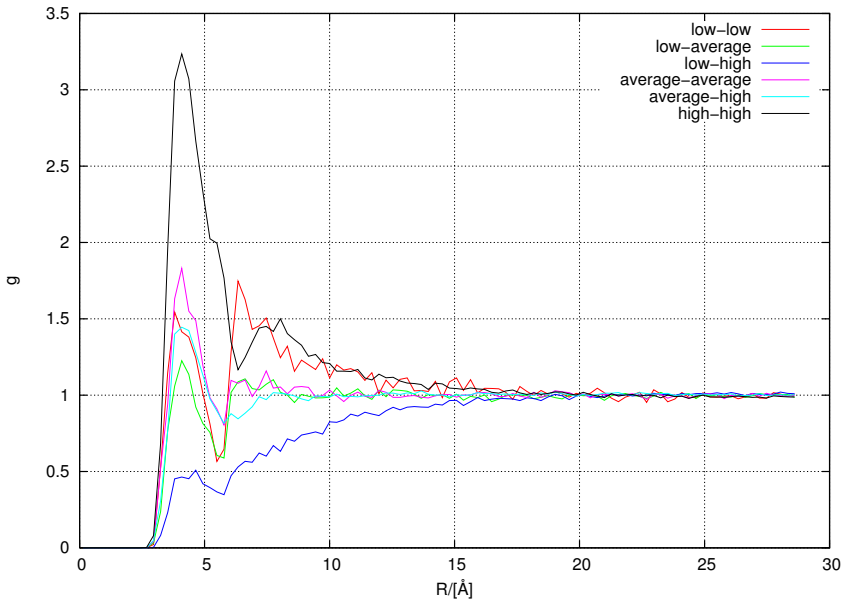


Figure 8.7.: Partial g-Function $r_{ck} = 6$ Å

The partial g-functions have a jump at 6 Å, this is a side effect of the spherical cut-off. The high-high and low-low regions show a significant second coordination shell. This means that both regions are compact. The low-low maxima are shifted towards smaller distances. All average combinations are gas-like at distances above r_{ck}. The low-high curve is smooth with nearly no first coordination shell. Higher order shells do not exist. Both regions are well separated, with only a small overlap region. The curve is below the gas density up to 20 Å.

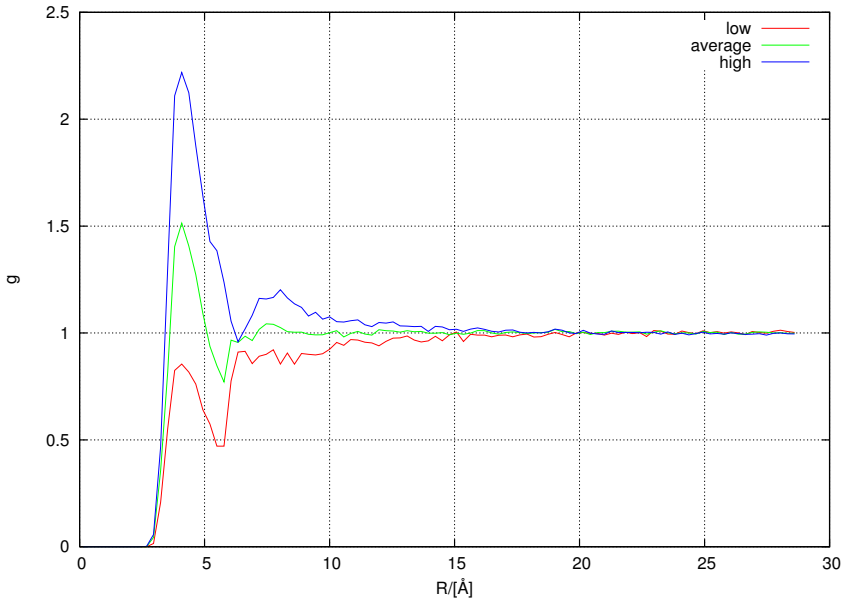


Figure 8.8.: Total g-Function $r_{ck} = 6\,\text{Å}$

The average region is surrounded by an ideal gas-like region. The high region approaches the ideal gas distribution from above, the central molecule is surrounded by more molecules up to 15 Å. The opposite is true for the low density region.

10 Å

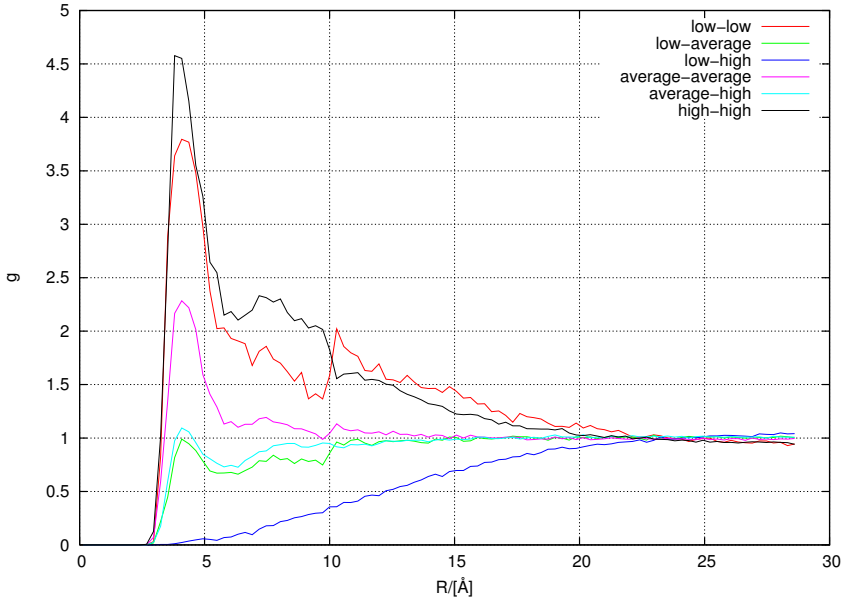


Figure 8.9.: Partial g-Function $r_{ck} = 10\,\text{Å}$

No structure is visible at all in the low-high curve. The average values get closer to the normal g-function. Three coordination shells can be seen in the high-high graph, whereas the low-low graph has no significant second order shell. The first maxima of all graphs are approximately at the same distance. At 23 Å are all graphs gas-like, but afterwards they diverge again in the opposite direction. The low and high regions are well separated.

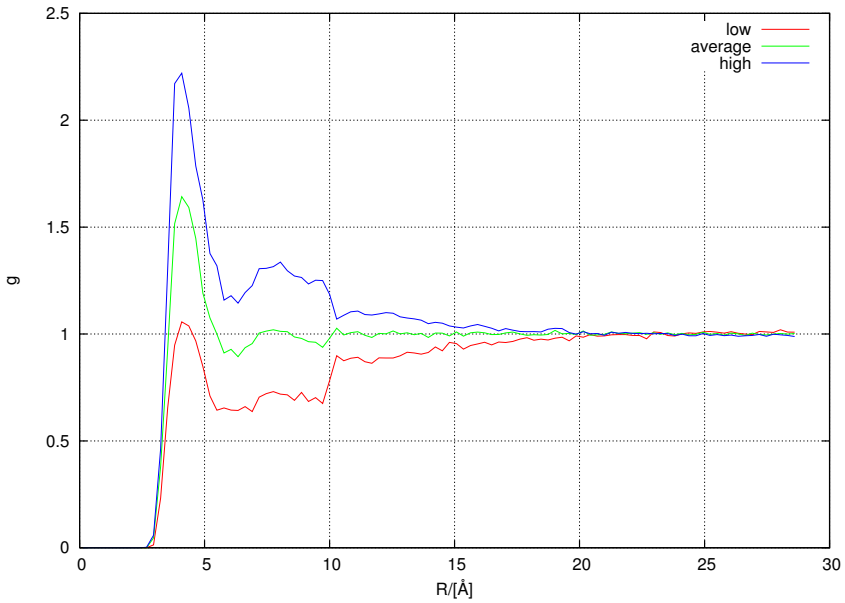


Figure 8.10.: Total g-Function $r_{ck} = 10\,\text{Å}$

There is not much change in the total g-function between 6 Å and 10 Å. It is now obvious that the low density molecules only have one coordination shell with a gas-like maximum.

20 Å

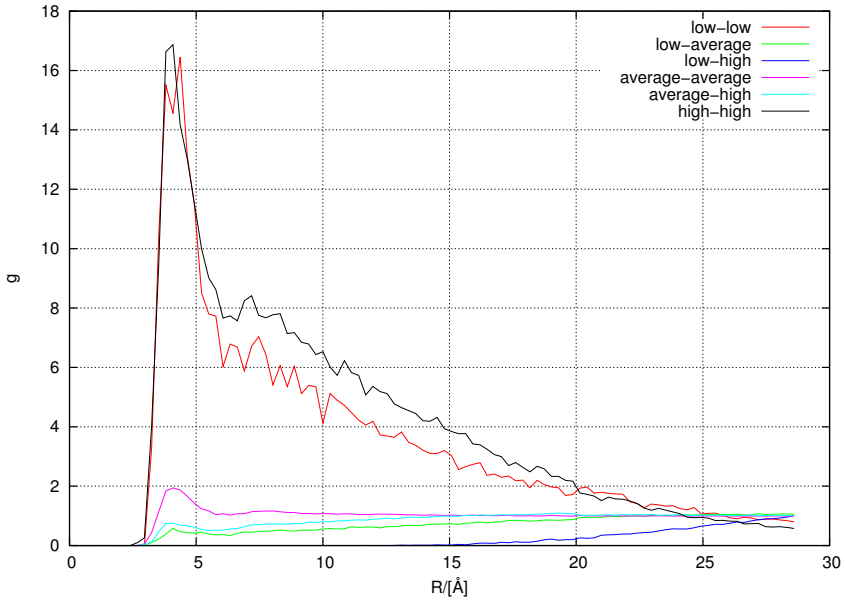


Figure 8.11.: Partial g-Function $r_{ck} = 20\,\text{Å}$

The differences between the partial g-functions get smaller.

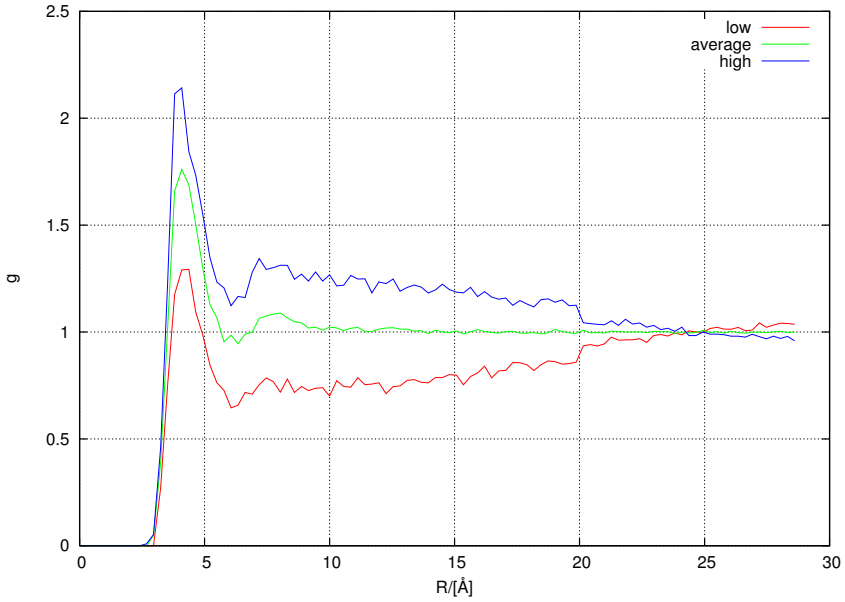


Figure 8.12.: Total g-Function $r_{ck} = 20\,\text{Å}$

The total g-functions get more similar, too. The box length of 57.7 Å limits the maximal r_{ck} value to about 29 Å, where 20 Å is already close to this value. Therefore the low and high

density regions disappear at larger r_{ck} values.

8.2. Time Auto-Correlation Function

So far static average values have been studied. To get the lifetime of the local density regions, a time correlation function[46, 85] is used

$$\begin{aligned} \delta\rho_{ck}(t) &= \rho_{ck}(t) - \bar{\rho}_{ck} && (8.10)\\ \bar{\rho}_{ck} &= \sum_{\rho_{ck}} \rho_{ck}\, \wp(\rho_{ck}) && (8.11)\\ C_p(t) &= \frac{\langle \delta\rho_{ck}(t)\, \delta\rho_{ck}(0)\rangle}{\left\langle [\delta\rho_{ck}(0)]^2 \right\rangle} && (8.12)\\ \tau_\rho &= \int_0^\infty \mathrm{d}t\, C_p(t) && (8.13) \end{aligned}$$

where $\delta\rho_{ck}$ is the molecular local density fluctuation variable and ρ_{ck} the local density in a sphere of a given radius. τ_ρ is called the lifetime or local density reorganization time, $\wp(\rho_{ck})$ is the probability of a local density.

In addition, the following local density averages are calculated:

Average	**Description**
$\langle \rho_{ck} \rangle$	Average local density
$\left\langle \frac{\rho_{ck}}{\rho_0} \right\rangle$	Average relative local density enhancement
$\langle \rho_{ck} - \rho_0 \rangle$	Average local density enhancement
$\left\langle \frac{\rho_{ck}-\rho_0}{\rho_0} \right\rangle$	Average relative local density difference

Where ρ_0 is the bulk density.

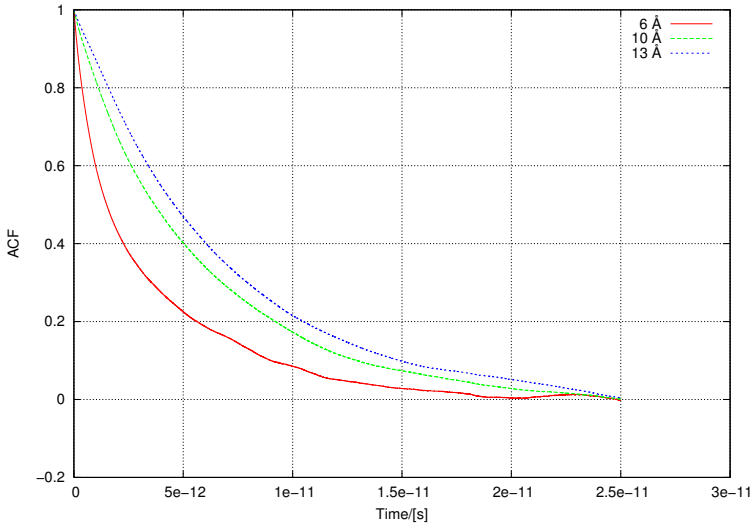


Figure 8.13.: CO_2 BBV Local Density Correlation Function at 300 K and 11500 mol m^{-3}.

Smaller density regions have a shorter lifetime due to the larger fluctuations.

All values of an isotherm at 300 K are shown in the following figure:

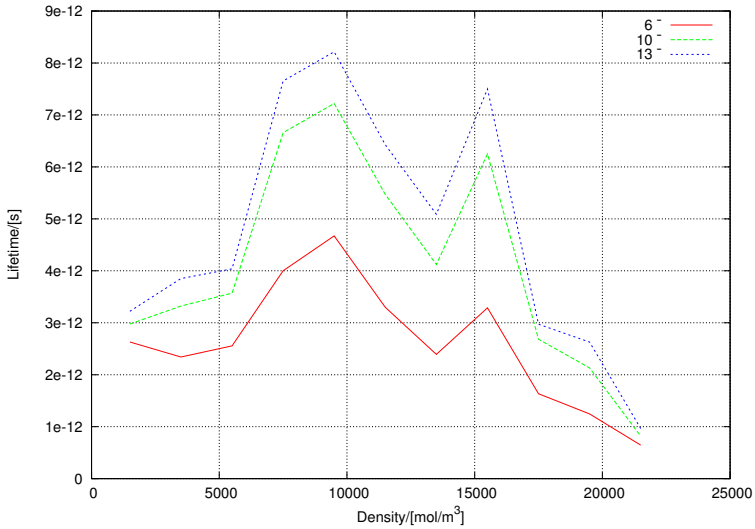


Figure 8.14.: CO_2 BBV Local Density Lifetime at 300 K.

Maddox *et al.* found a maximum in the lifetime at the critical density. This phenomenon is called critical slowing down. The density regions get more stable and have a larger influence on macroscopic properties. In figure 8.14 two maxima and a local minimum in the critical density region are visible. Either the simulation was in the two-phase region or the system is too small to sample large fluctuations. Indeed, the two maxima are close to the vapor-liquid densities recently obtained by Gibbs ensemble molecular dynamics simulations (see histogram 7.22 on

page 96). An isotherm at a 15 % higher temperature should give a maximum at the critical density.

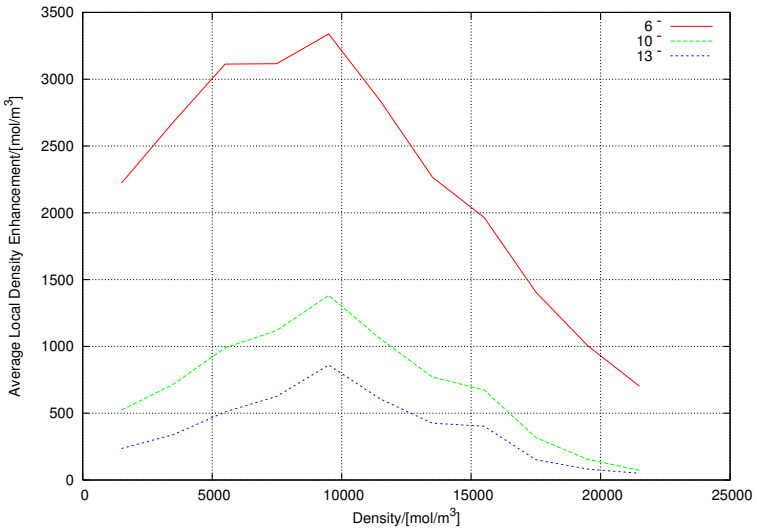


Figure 8.15.: CO_2 BBV Average Local Density Enhancement at 300 K.

The maximum in the local density enhancement is below the critical density. The same behavior was found by Tucker *et al.*[84, 47] and in experiments (see section 8.3). The enhancement is larger at smaller cut-off radii.

A plot of the density probabilities[47] is shown in the next figure. The local density distribution is broadest near the critical point. This observation was also seen while using the Gibbs ensemble molecular dynamics method above the critical temperature.

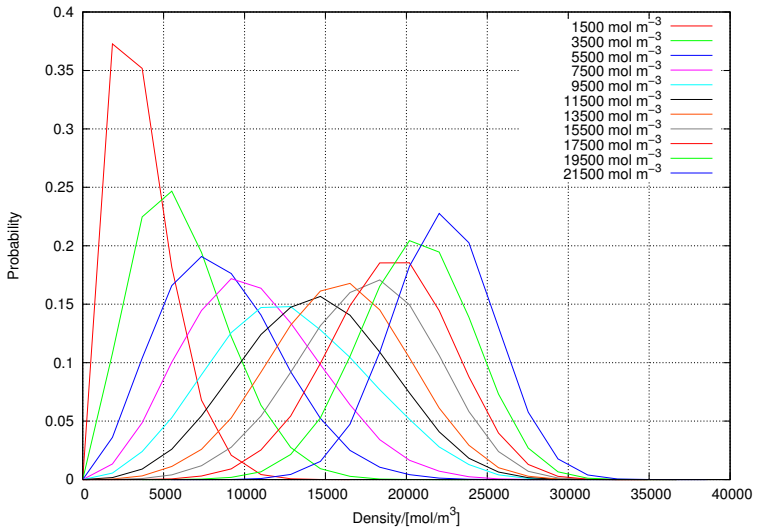


Figure 8.16.: CO_2 BBV Local Density Probability at 300 K and $r_{ck} = 6$ Å.

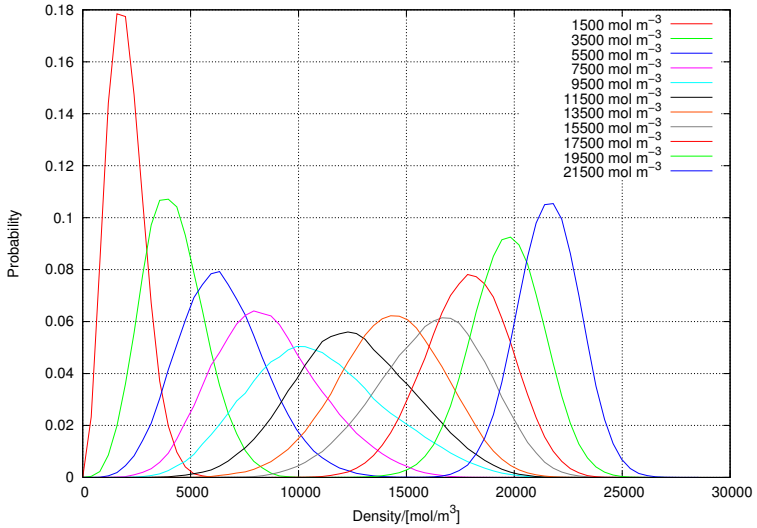


Figure 8.17.: CO_2 BBV Local Density Probability at 300 K and $r_{ck} = 10$ Å.

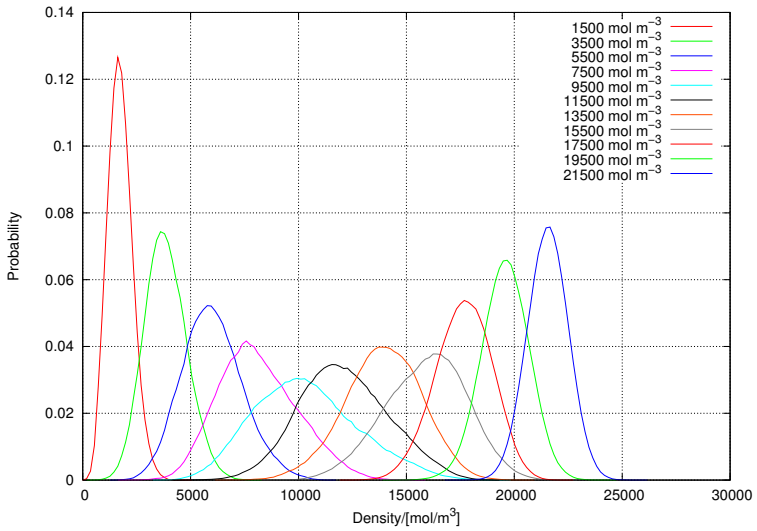


Figure 8.18.: CO_2 BBV Local Density Probability at 300 K and $r_{ck} = 13$ Å.

The broadest distribution is between 9500 and 11500 mol m^{-3}.

Local density specific radial distribution functions[26] were also calculated. Their shape is similar to the total g-functions shown before. The resolution is higher since the sampling is done at each correlation step.

8.3. Experimental Data

8.3.1. Vibrational Lifetime

The vibrational lifetime of $W(CO)_6$ in supercritical CO_2 shows a plateau on a near critical isotherm 1 % above the critical temperature[89, 90]. In the vicinity of the critical point, the observables are independent of the density. At a temperature 17 K higher a steady decay was measured.

Goodyear and Tucker added a rigid diatomic molecule to their 2-D Lennard-Jones fluid and calculated the vibrational population relaxation time by a time auto-correlation function[28]. The relaxation time behavior was equal to the plateau obtained in the experimental data.

8.3.2. Raman Spectral Shift

The vibrational frequency shift of Raman spectra shows a plateau region near the critical point[70].

8.3.3. X-Ray Scattering

Nishikawa *et al.* used small-angle X-ray scattering experiments to calculate density fluctuations and the correlation length[56]. CO_2, CF_3H and C_2H_4 measurements gave the same order of fluctuations. CO_2 has a strong quadrupole moment, C_2H_4 a weak quadrupole moment and CF_3H a strong dipole. The critical behavior is independent of the substance. The maxima are at a density slightly below the critical density. At a temperature 2 % above the critical temperature, the measured correlation length was 20 Å.

9. Summary and Outlook

Gibbs Ensemble

A new algorithm for the Gibbs ensemble molecular dynamics method was constructed (see section 6.2 on page 56). The algorithm is based on the Nosé-Hoover non-Hamiltonian equations of motion. It was shown that states defined by the Gibbs ensemble partition function are sampled (see section 6.2.4 on page 70 and previous pages).

The molecular dynamics algorithm is more complicated than the Monte Carlo method. However, the complexity does not increase from atoms to larger molecules, whereas the Monte Carlo method has to be extended to sample additional degrees of freedom. Further advantages are that time averages are obtained, time correlation functions can be calculated and the single step transfer could be replaced by an exchange trajectory.

An accurate *ab initio* CO_2 pair potential was used to validate the algorithm. The critical values were in good agreement with the experimental critical point. The molecular results will be the first published data of the Gibbs ensemble molecular dynamics method with one exchange molecule. It was shown that the exchange length is proportional to the density of the liquid phase. Therefore, the same method could even be used for solid phases.

The new Gibbs ensemble molecular dynamics algorithm could be easily extended to larger molecules by extending the exchange algorithm or to smaller systems like Car-Parrinello molecular dynamics by adding Nosé-Hoover chains. The non-Hamiltonian approach gives a high flexibility for the construction of equations of motion in the future.

It would be interesting to check the algorithm with different potentials under various conditions such as large molecules, solid phases or the triple point. It should be possible to get the triple point by using three subsystems for the liquid, vapor and solid phases.

Critical Properties

The properties of supercritical CO_2 in the vicinity of the critical point were studied by different methods. Several critical phenomena known from experiments could be reproduced.

To get a better insight into the critical region and the correlation length in the vicinity of the critical point, the calculation of more isotherms is necessary. So far, the microcanonical ensemble was used to calculate the time correlation. The results could be improved by using a canonical ensemble.

Appendix

A. Gibbs Ensemble SimMol User Information

A.1. Compilation

The following shell command builds the simulation program `sm-co2rig-verlet_ge-Xeon`[11]. `CO2RIG` can be exchanged by `CO2FLX` or `CO2SAPT` to use flexible molecules or the SAPT potential.

```
./build SIM_MOL CO2RIG VERLET_GE NO_CORR
```

The correlation manager cannot be used in this program version because the box length scaling is not handled and the number of molecules changes. Molecules disappear and are newly created.

The program uses the LAM[45][1] implementation of the Message Passing Interface (MPI). The include file `mpif_lam.h` has to match the header file of the LAM runtime installed on the system, otherwise the program will crash.

It will be shown next how LAM is compiled to use Intel's Fortran compiler. This is only needed on the machine where the program is compiled. On the target systems the LAM runtime is sufficient. The example below shows how to install LAM 7.1.1.

On a Debian system one gets the source code of the package `lam4`: `lam_7.1.orig.tar.gz` and the patches `lam_7.1-1.diff.gz`. The tar file `lam_7.1.orig.tar.gz` has to be extracted. The directory `lam-7.1` was created. Change to this directory and run "`patch -p1 < ../lam_7.1-1.diff`" to apply the patches. Then use these commands to configure and install LAM:

```
export CXX=g++
./configure --prefix=/usr --with-fc=ifc --with-f90=ifc
make
make install
```

Synchronize the header file `lam-7.1/share/include/mpif.h` with `mpif_lam.h` of the SimMol program.

[1] Local Area Multicomputer

A.2. Execution

The input files for each simulation box (subsystem) have to be created by the input program. How to build the simulation program was described in the previous section.

The LAM runtime is started with `lamboot`, if it is not yet running. In the file `apprun`, a command to execute a single process is written on each line. The example shows which commands have to be used to start a two-phase simulation:

```
sm-co2rig-verlet_ge-Xeon CO2_220_box1 output1.txt
sm-co2rig-verlet_ge-Xeon CO2_220_box2 output2.txt
```

The first parameter is the usual stem name of the input files, the second parameter is the file name where the standard output is redirected to. If the standard output is not redirected, then LAM collects all outputs and redirects them to the standard output of the start program, which is not useful at all.

The LAM start command, in the `run` file, looks like this:

```
mpirun apprun > output.txt
```

The `output.txt` should be empty if the above `apprun` is used.

The whole simulation, using two processes, is started with:

```
psub run
```

A.3. .EXT file

The extended ensemble status is written to the file "*stem name*.EXT" for each simulation box. The file contains six columns with the following values:

Column	Value
1	Step number.
2	Density in $\frac{mol}{m^3}$.
3	Volume in $Å^3$.
4	Number of molecules.
5	Molecular pressure in k Pa.
6	Conserved energy in $\frac{J}{mol}$.

Table A.1.: SimMol .EXT File.

The values are written in intervals of IPRINT steps. Columns two to six are the average values of these steps.

A.4. .DHG file

The .DHG file contains the density histogram[24, p. 216] of this run. If a run is continued, then the histogram values are reset.

The file has two columns:

Column	Value
1	Density in $\frac{mol}{m^3}$.
2	Number of occurrences.

Table A.2.: SimMol .XCH File.

The values are written in intervals of SAVEHI steps to a new file "*stem name*step.DHG". It is stored in the same intervals as the other histograms.

A.5. .XCH file

The exchange file contains a log of all successful and failed particle exchanges. Below is an example:

```
Particle exchange log file
--------------------------
    101 source: 1 dest: 0
   2055 exchange was successful (1955 steps)
   2150 source: 1 dest: 0
   3969 exchange failed (1820 steps)
   3975 source: 0 dest: 1
   5018 exchange failed (1044 steps)
   5030 source: 1 dest: 0
   7259 exchange failed (2230 steps)
   7342 source: 0 dest: 1
   8791 exchange failed (1450 steps)
   8847 source: 1 dest: 0
  11181 exchange was successful (2335 steps)
  11207 source: 0 dest: 1
  11852 exchange was successful (646 steps)
  11936 source: 1 dest: 0
  13525 exchange was successful (1590 steps)
  13595 source: 0 dest: 1
  14743 exchange was successful (1149 steps)
```

In the first column is the time step and after that a human readable output. In step 101 was a fractional particle created in box 0, the exchange from box 1 to box 0 was successful in 1955 time steps. Note, that the box indices start from zero.

A.6. Povray

The program can create Povray files at the end of successful exchanges. The following global constants of `verlet_ge.f90` control this:

Constant	Value
GE_POV_USE	Enable the Povray output.
GE_POV_INT	Interval of Povray outputs in steps of successful exchanges.

Table A.3.: Povray SimMol Constants.

The files are stored as "`step number[s/d].pov`". With s for source and d for the destination box.

The source code in the file `povray.f90` generates the Povray input file. The Povray generation code is based on values from OpenBabel[75]. The unit cell is stored in the Povray file. All molecules which pierce the unit cell are repeated at positions where some atoms lie in the box. The fractional molecule is shown in blue.

The `.pov` file can be rendered with Povray to produce a pixel image:

```
povray +W500 +H500 +Q9 +A 120493d.pov
```

A movie showing the rotating box can be produced with these commands:

```
povray +W500 +H500 +KFI0 +KFF100 120493d.pov
mencoder "mf://*.png" -mf w=500:h=500:fps=20:type=png -ovc divx4 -divx4opts q=5:br=8000 -o output.avi
```

A.7. XMol Trajectories

Trajectories of the molecules during particle exchanges can be written to XMol files. The creation is controlled by:

Constant	Value
GE_XMOL_USE	Enable the trajectory creation.
GE_XMOL_SAVE	Interval between snapshots in time step units.

Table A.4.: XMol SimMol Constants.

The trajectories of failed exchanges are removed. Nevertheless, the trajectory files get very large. This option should only be used for debugging reasons.

The XMol files are stored as "exchange*number*[**s**/**d**].xyz". With s for source and d for the destination box.

Jmol[74] is recommended to visualize the xyz movies.

A.8. Final Output

At the end of the simulation, statistics of the production cycle are written to the screen.

```
 Conserved energy:    7107.7642   0.2129E+01
 Conserved energy shifted:    7119.9023   0.2129E+01

 Start density in mol/m^3: 20360.8362
 Average density in mol/m^3: 21392.1571      1.5546
 End density in mol/m^3: 22188.6165

 Atomic pressure in bar:       43.5765      0.2473
 Molecular pressure in bar:    42.9483      0.2473
 Long-range pressure in bar:    -4.2539      0.0012
 GE atomic pressure in bar:    43.5246
 GE molecular pressure in bar:    43.1627

 Atomic pressure in bar (w/o LR):       47.8304      0.2476
 Molecular pressure in bar (w/o LR):    47.2022      0.2476

NPT ensemble values:
 Start volume in m^3: 0.1702E-25
 Volume in m^3: 0.1643E-25  0.1235E-29

Gibbs ensemble values:
 Total volume in m^3:   0.8834E-25   0.0000E+00
 Molecules in box:        211.10         0.00
 Molecules in box now:   211

Ensemble statistics:
 Min molecular pressure:   -625.3691 bar (step: 907575)
 Max molecular pressure:    909.6887 bar (step: 907524)

 Min thermo. temperature: 227.38496 K (step: 518740)
 Max thermo. temperature: 318.46111 K (step: 782173)

Gibbs ensemble algorithm statistics:
 Successful exchanges: 176
 Failed exchanges: 3002
```

```
 Average exchange length:   1242.9943
 Average exch. failed length:    207.2765

 Number of times source box: 1543
 Number of times dest box: 1635

 Min rcut [Ang]:     12.0698
 Max rcut [Ang]:     12.8657

 Rcut [Ang]:     12.5335
 Rcut*:      0.5000

Berechnung thermodynamischer Grössen

Durchschnitte der 500000 Schritte

 Temperature:  270.00 K
 Pressure:      43.58 bar
 Pressure:       4.36 MPa
```

B. Technical Program Changes

In this chapter, several changes in the SimMol program since the beginning of my PhD will be documented. The detailed documentation of the original microcanonical program version can be found in my diploma work[11].

B.1. Input program

B.1.1. Lattice orientation

If a random orientation of the molecules on their lattice position was chosen, it is possible that atoms of one molecule are close to atoms of an immediate neighbor molecule. Therefore the initial energy of the system is high. The rotational energy rapidly increases in the first simulation steps and it takes more velocity rescaling steps to reach the desired equilibrium temperature.

A simple solution to this problem is to rotate the molecules which are too close to optimize the distances.

The source code is in the subroutine `ROTATE_MOLECULES` of `Main.f`. At the beginning, the molecules are oriented according to their Euler angles. If the logical variable `OPTIMIZE` is true then the algorithm starts:

```
SUCCESS = .FALSE.
DO TRIES = 1,MAXTRIES
   CALL CALL GET_MIN_INTER(DIST,...,NR)
   COUNT = 0
   DO I = 1,NR
      IF (DIST(I).LT.MAXDIST) THEN
        COUNT = COUNT+1
        CALL ROTATE_MOLECULE(...)
      END IF
   END DO
   IF (COUNT.EQ.0) THEN
      SUCCESS = .TRUE.
      EXIT
   END IF
END DO
```

The default values are:

OPTIMIZE	.TRUE.
MAXDIST	200.D−12
NR	40
MAXTRIES	100

Example output of a CO_2 input:

```
Some molecules are too near, starting optimization:
      24 molecules randomly rotated!!!
       1 molecules randomly rotated!!!
Done in   3 steps.
Shortest intermolecular distances:
        200.5215 pm between particle   63 (O , 1) and   91 (O , 1)
        200.8469 pm between particle  841 (O , 1) and 1023 (O , 1)
        201.5971 pm between particle  798 (O , 1) and  801 (O , 1)
        201.9747 pm between particle  316 (O , 1) and  323 (O , 1)
        202.2353 pm between particle   71 (O , 1) and   78 (O , 1)
```

For pure systems, the number of rotations is small. This algorithm was added for mixtures with molecules of different size: *e.g.* Water/DMSO. If two or more large molecules are close, then it takes many steps to optimize the system:

```
Some molecules are too near, starting optimization:
      40 molecules randomly rotated!!!
      37 molecules randomly rotated!!!
      17 molecules randomly rotated!!!
      14 molecules randomly rotated!!!
       3 molecules randomly rotated!!!
       2 molecules randomly rotated!!!
Done in   7 steps.
Shortest intermolecular distances:
        200.3644 pm between particle 1178 (H , 1) and 1233 (Me, 2)
        200.8710 pm between particle  593 (O , 2) and  939 (Me, 2)
        201.7795 pm between particle  852 (Me, 2) and  855 (Me, 2)
        203.7184 pm between particle  730 (Me, 2) and  733 (Me, 2)
        204.0356 pm between particle  617 (Me, 2) and  628 (Me, 2)
```

B.1.2. Vibrations

In the original SimMol code the normal modes were defined by a force constant, displacement vectors and the equilibrium distance. A Gaussian distribution was used to assign a kinetic energy. The atoms were displaced to set the potential energy.

The problem of this method was that potential and kinetic energy are not related. Large fluctuations of the temperature were visible in the simulations. Because the energy exchange between the vibration and the other degrees of freedom is very slow, it was a constant effect.

In the new code the phase of each normal mode is a random number between zero and 2π. Each normal mode has the exact total energy with a different kinetic and potential energy. The input parameters specify the intramolecular force field given by linear and angular force constants. A normal mode analysis is implemented in the program. The code in `CALC_vib.f` (module `SMI_VIBCA`) is based on the program VIBCA[22]. It was modified to support dummy atoms.

The force field data is stored in the `VIBRATION` structure:

Field	Value	Description
`NAME`	CHARACTER(15)	Name of the vibration.
`MTAG`	INTEGER	Molecular species to which the vibration belongs.
`ICC1`	INTEGER	First force field definition.
`ICC2`	INTEGER	Second force field definition.
`FLAG`	INTEGER	User defined data.
`LIST(12,2)`	INTEGER	Coordinates of the force field.
`FORCE_CONSTANT`	REAL(8)	Spring constant in mdyn Å^{-1} (stretch), mdyn rad^{-1} (bend) or mdyn Å rad^{-2}.
`FORCE_CONSTANT_SI`	REAL(8)	Spring constant in N m^{-1} or N rad^{-1}.
`MODES`	TYPE(NORMALMODE)	Pointer to the calculated normal mode data.

Table B.1.: VIBCA VIBRATION Structure.

1 dyn is equal to 10^{-5} N.

`NORMALMODE` is defined as:

Field	Value	Description
`OMEGA`	REAL(8)	Frequency in s^{-1}.
`MAXDISPLX(:)`	REAL(8)	Maximum displacement [m] for the simulation temperature. Only for real atoms.
`MAXDISPLY(:)`	REAL(8)	
`MAXDISPLZ(:)`	REAL(8)	
`NEXT`	TYPE(NORMALMODE)	Pointer to the next item.

Table B.2.: VIBCA NORMALMODE Structure.

The following force field values are possible (variables `ICC1` and `ICC2`):

ICC_NONE No force field.

ICC_STRETCH Linear stretch between LIST(1,I) and LIST(2,I).

ICC_BEND Bending of the angle LIST(1,I)-LIST(2,I)-LIST(3,I).

ICC_OUTOFPLANE Four coplanar atoms. The atom LIST(1,I) is bending out of the plane LIST(2,I)-LIST(4,I)-LIST(3,I).

ICC_LINEARBEND Bending of the collinear atoms LIST(1,I)-LIST(2,I)-LIST(3,I) on the plane given by atom LIST(4,I). If atom LIST(4,I) is not an index of a real atom then both orthogonal positions are automatically calculated[1].

ICC_4ATORSION A torsion of the atoms LIST(1,I)-LIST(2,I)-LIST(3,I)-LIST(4,I).

ICC_DUMMY Dummy site LIST(1,I) in the middle of the atoms LIST(2,I) and LIST(3,I).

The example below is the force field of the 5-center CO_2 potential:

Variable	Value	Description
NAME	"O1-C2"	
MTAG	1	
ICC1	ICC_STRETCH	Linear stretching.
ICC2	ICC_NONE	Not used.
FLAG	0	
LIST(1,1)	1	First oxygen.
LIST(2,1)	2	Carbon.
FORCE_CONSTANT_SI	CORE(1)	
FORCE_CONSTANT	CORE(1)/M2A*N2MDYN	
NAME	"C2-O3"	
MTAG	1	
ICC1	ICC_STRETCH	Linear stretching.
ICC2	ICC_NONE	Not used.
FLAG	0	
LIST(1,1)	2	Carbon.
LIST(2,1)	3	Second oxygen.
FORCE_CONSTANT_SI	CORE(1)	
FORCE_CONSTANT	CORE(1)/M2A*N2MDYN	
NAME	"O1-O3"	
MTAG	1	
ICC1	ICC_STRETCH	Linear stretching.
ICC2	ICC_NONE	Not used.
FLAG	0	
LIST(1,1)	1	First oxygen.
LIST(2,1)	3	Second oxygen.
FORCE_CONSTANT_SI	CORE(2)	

[1] See CO_2 example output.

Variable	Value	Description
FORCE_CONSTANT	CORE(2)/M2A*N2MDYN	
NAME	"O1-C2-O3"	
MTAG	1	
ICC1	ICC_LINEARBEND	Linear molecule with out of plane bend.
ICC2	ICC_NONE	Not used.
FLAG	0	
LIST(1,1)	1	First oxygen.
LIST(2,1)	2	Carbon.
LIST(3,1)	3	Second oxygen.
LIST(4,1)	4	Dummy site 1 (orthogonal).
FORCE_CONSTANT_SI	COANG(1)	
FORCE_CONSTANT	COANG(1)*M2A*N2MDYN	
NAME	"Dummy 1"	
MTAG	1	
ICC1	ICC_DUMMY	Dummy site.
ICC2	ICC_NONE	Not used.
FLAG	0	
LIST(1,1)	4	This dummy site.
LIST(2,1)	1	First oxygen.
LIST(3,1)	2	Carbon.
NAME	"Dummy 2"	
MTAG	1	
ICC1	ICC_DUMMY	Dummy site.
ICC2	ICC_NONE	Not used.
FLAG	0	
LIST(1,1)	5	This dummy site.
LIST(2,1)	2	Carbon.
LIST(3,1)	3	Second oxygen.

Table B.3.: CO_2 Force Field.

Output of the VIBCA program code:

```
Entering the CALC_VIBRATION procedure

CARTESIAN COORDINATES (Angstr):
-------------------------------
ATOM NO.         X             Y             Z       MASS (amu)

 1          -1.161000      0.000000      0.000000     16.000000
 2           0.000000      0.000000      0.000000     12.000000
 3           1.161000      0.000000      0.000000     16.000000
 4           0.000000      0.707107      0.707107      0.000000
 5           0.000000     -0.707107      0.707107      0.000000
```

```
ROTATION MATRIX TO PRINCIPAL COORDINATES:
-----------------------------------------
               XOLD                YOLD                ZOLD

a            1.000000            0.000000            0.000000
b            0.000000            1.000000            0.000000
c            0.000000            0.000000            1.000000
determinant= 1.000000

PRINCIPAL COORDINATES (Angstr):
-------------------------------
ATOM NO.         a             b             c       MASS (amu)

 1          -1.161000      0.000000      0.000000     16.000000
 2           0.000000      0.000000      0.000000     12.000000
 3           1.161000      0.000000      0.000000     16.000000
 4           0.000000      0.707107      0.707107      0.000000
 5           0.000000     -0.707107      0.707107      0.000000

------------------------------------------------------------------------------
SPECIFICATION OF MOLECULAR FORCE FIELD (in mdyn/A etc. units):
------------------------------------------------------------------------------
   FORCE CONSTANTS                        COORDINATES
TYPE  INDEX    VALUE                ATOMS              ATOMS     NAME

i,i    1    15.066000      STR.  1  2  0  0         0  0  0  0   O1-C2
i,i    2    15.066000      STR.  2  3  0  0         0  0  0  0   C2-O3
i,i    3     0.865150      STR.  1  3  0  0         0  0  0  0   O1-O3
i,i    4     0.718990      LIBE  1  2  3  4         0  0  0  0   O1-C2-O3

FREQUENCIES (cm-1) AND CARTESIAN COORDINATES OF NORMAL MODES
-----------------------------------------------------------
FREQUENCY          MASS-WEIGHTED CARTESIANS          UNWEIGHTED CARTESIANS
          ATOM      a         b         c            a         b         c
Mode  1:
 644.1565   1    0.000000 -0.261116 -0.261116     0.000000 -0.234261 -0.234261
            2    0.000000  0.603023  0.603023     0.000000  0.624695  0.624695
            3    0.000000 -0.261116 -0.261116     0.000000 -0.234261 -0.234261
            Kin. T, av. over phases =     262.4133 K
Mode  2:
1334.8064   1   -0.707107  0.000000  0.000000    -0.707107  0.000000  0.000000
            2    0.000000  0.000000  0.000000     0.000000  0.000000  0.000000
            3    0.707107  0.000000  0.000000     0.707107  0.000000  0.000000
            Kin. T, av. over phases =     276.8412 K
Mode  3:
2420.7289   1   -0.369274  0.000000  0.000000    -0.331295  0.000000  0.000000
            2    0.852803  0.000000  0.000000     0.883452  0.000000  0.000000
            3   -0.369274  0.000000  0.000000    -0.331295  0.000000  0.000000
            Kin. T, av. over phases =     309.1018 K

All modes: Kin. T, av. over phases and mod. =     282.7854 K


------------------------------------------------------------------------------
B MATRIX (NINTx3N)   a. 1   b. 1   c. 1   a. 2   b. 2   c. 2   a. 3   b. 3   c. 3

STR.  1, 2, 0, 0  -1.0000 0.0000 0.0000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000
STR.  2, 3, 0, 0   0.0000 0.0000 0.0000-1.0000 0.0000 0.0000 1.0000 0.0000 0.0000
STR.  1, 3, 0, 0  -1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000 0.0000 0.0000
LIBE  1, 2, 3, 4   0.0000 0.6090 0.6090 0.0000-1.2181-1.2181 0.0000 0.6090 0.6090

------------------------------------------------------------------------------
G MATRIX = B M**-1 B' (NINTxNINT):
------------------------------------------------------------------------------
 1  0.145833 -0.083333  0.062500  0.000000
 2 -0.083333  0.145833  0.062500  0.000000
```

```
 3  0.062500  0.062500  0.125000  0.000000
 4  0.000000  0.000000  0.000000  0.340030

--------------------------------------------------------------------------------
L MATRIX = B M**-1/2 l (NINTxNMODES, normalized to G = LL'):
--------------------------------------------------------------------------------
 1  0.000000  0.176777  0.338502
 2  0.000000  0.176777 -0.338502
 3  0.000000  0.353553  0.000000
 4 -0.583121  0.000000  0.000000

--------------------------------------------------------------------------------
POTENTIAL ENERGY DISTRIBUTION (diagonal force constant contributions only):
--------------------------------------------------------------------------------
                      1   2   3
 Normal mode ->     644.2  2420.7
                       1334.8

STR.  1, 2, 0, 0     0  45  50
STR.  2, 3, 0, 0     0  45  50
STR.  1, 3, 0, 0     0  10   0
LIBE  1, 2, 3, 4   100   0   0

Leaving the CALC_VIBRATION procedure
```

Due to the random phases, the kinetic temperature of the initial oscillations is not equal to the system temperature. But the average temperature will be equal to the average temperatures of the translation and rotation, giving the expected overall fluctuations.

B.1.3. BBV Input

The BBV potential uses the same 5-center site-site model, potential and damping function as the 8s potential. Nothing was changed there.

B.1.4. SAPT-s Input

The SAPT[12] potential is defined in section 7.1.2 on page 76.

$CO2_SAPT-s

Written by `CO2SAPT0`.

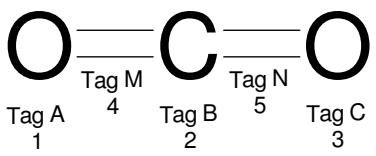


Figure B.1.: CO_2 Site Tags

Variable	Type	Count	Description
TAGA	'␣TAGA:', T40, I10	1	First oxygen atom.
TAGB	'␣TAGB:', T40, I10	1	Carbon atom.
TAGC	'␣TAGC:', T40, I10	1	Second oxygen atom.
TAGM	'␣TAGM:', T40, I10	1	First dummy site.
TAGN	'␣TAGN:', T40, I10	1	Second dummy site.
NOPAIR	'␣NOPAIR:', T40, I10	1	Number of atom-atom interactions.

NOPAIR times:

Variable	Type	Count	Description
ATOMA(I), ATOMB(I)	'␣ATOMA,ATOMB:', T40, 2I10	1	Interaction between atom A and B.
ALPHA(I)	'␣ALPHA:', T40, G25.16	1	Parameter ALPHA.
BETA(I)	'␣BETA:', T40, G25.16	1	Parameter BETA.
DELTA1(I)	'␣DELTA1:', T40, G25.16	1	Parameter DELTA1.
C6(I)	'␣C6:', T40, G25.16	1	Parameter C6.
DELTA6(I)	'␣DELTA6:', T40, G25.16	1	Parameter DELTA6.
C8(I)	'␣C8:', T40, G25.16	1	Parameter C8.
DELTA8(I)	'␣DELTA8:', T40, G25.16	1	Parameter DELTA8.

Variable	Type	Count	Description
RCUT	'␣RCUT:', T40, G25.16	1	Cut-off radius [m].
EQCO	'␣EQCO:', T40, G25.16	1	C-O equilibrium distance.
EQCM	'␣EQCM:', T40, G25.16	1	C-M equilibrium distance.

B.1.5. Verlet_GE Input

The Verlet algorithm for extended ensembles supports these ensembles: NVE, NVT, NPE, NPT and Gibbs.

Subroutine: SIMALG

```
 Input the simulation algorithm used.
 The available algorithms are:
  LEAPFROG                  $LEAPFROG
  VERLET                    $VERLET
  VERLETGE                  $VERLETGE
$Simulation Algorithm:
```

The simulation algorithm.

Variable set: ALGOR

Verlet-Gibbs ensemble

Subroutine: VERLETGEI

```
Type of ensemble to use:
(1) NVE
(2) NVT
(3) NPT/GE
```

Variables set: ET ('NVE', 'NVT' or 'NPT')

NVE uses the normal microcanonical Verlet algorithm. NVT is reserved for an optimized NVT algorithm. If a NVT ensemble is needed, one has to select NPT and set the barostat mass to zero. This deactivates the volume scaling.

For Gibbs ensemble simulations, the simulation start files have to be created for each box (phase).

If NVE was selected then the input ends here. For the NPT case additional parameters are read:

```
Thermostatization during the equilibration:
(1) Velocity scaling
(2) Nose-Hoover thermostat
(3) Hoover NPT
```

Variables set: TC ('NVE', 'NVT' or 'NPT')

Defines how the temperature is controlled during the equilibration. Option 1 uses the normal velocity rescaling.

```
Mass of the thermostat [kg m^2]
```

Variables set: VG_Q

Reads the mass of the thermostat in SI units. The thermostat temperature and the total degrees of freedom are known from the earlier input. The thermostat frequency is a function of the system size (see equation (5.79) on page 38).

```
Mass of the barostat [kg m^2] (0: none)
```

Variables set: VG_QP

Reads the barostat mass in SI units. Zero is reserved for no barostatization.

```
Barostat pressure [Pa]
```

Variables set: VG_P_FIX

Reads the external pressure for the NPT ensemble. For Gibbs ensemble simulations this value has no meaning.

```
Scale rcut with box length? (j/n)
```

Variables set: VG_SCALERC

The cut-off radius can be constant in meters or be scaled together with the size of the box (constant in reduced units). If the radius is scaled, then the cut-off value is always smaller than half the box length. Otherwise with a fixed radius it is possible that the radius gets larger than the maximum allowed value. In this case the simulation will be stopped. The input program writes a warning statement:

```
Warning: rcut is half the box's length!
   Volume scaling could stop the program.
```

```
Prepare for Gibbs ensemble? (j/n)
```

Variables set: USE_GE

Asks if the additional parameters for the Gibbs ensemble should be read. Without these parameters SimMol can only be used for single phase simulations in the NVT or NPT ensemble.

The Gibbs ensemble parameters are:

```
Mass of the exchange [kg m^2]
```

Variables set: VG_QE

Reads the exchange mass in SI units.

```
Fractional distance increase eta [m]:
```

Variables set: VG_FRAC_ETA

Reads the fractional distance increase parameter η.

```
Nu start position [0..1]:
```

Variables set: VG_NU_POS

The start position of ν in each exchange, usually zero.

```
Nu start velocity [0..1]:
```

Variables set: VG_NU_VEL

The mean value of the initial exchange velocity in units of $\Delta\,\nu$ per time step.

```
Start with exchanges at step:
```

Variables set: VG_EXCSTART

Number of time steps without exchanges after the equilibration.

```
Break between exchanges:
```

Variables set: VG_EXCBREAK

Maximum number of steps between two exchanges.

The Gibbs ensemble input ends here. The next parameters are again for all ensembles.

```
Write density histogram? (j/n)
```

Variables set: USE_DHG

If selected, then during the simulation `.DHG` files are created. The following values define the range and number of intervals of the histogram:

```
Start density:
```

Variables set: DHG_START

Start density in $\frac{mol}{m^3}$.

```
End density:
```

Variables set: DHG_END

End density in $\frac{mol}{m^3}$.

```
Intervals:
```

Variables set: DHG_INTS

Number of intervals.

B.1.6. Verlet_GE Output File Format

The initial values for the simulation program are written to the .PAR file. The Verlet-Gibbs ensemble section starts with the keyword:

$VERLETGE

Written by VERLETGEO.

Variable	Type	Count	Description
ET	'␣ENSEMBLE:', T40, A	1	Equilibration ensemble.

For The NVE ensemble:

Variable	Type	Count	Description
ET	'␣EQ.␣TEMP␣CONTROL:', T40, A	1	Equilibration temperature control.

For all other ensembles:

Variable	Type	Count	Description
TC	'␣EQ.␣TEMP␣CONTROL:', T40, A	1	Equilibration temperature control.
VG_Q	'␣NOSE␣MASS:', T40, G25.16	1	Thermostat mass.
VG_DFREE	'␣NOSE␣DEGREES␣OF␣FREEDOM:', T40, G25.16	1	Total degrees of freedom.
VG_T_FIX	'␣NOSE␣TEMPERATURE:', T40, G25.16	1	Thermostat temperature.
VG_QP	'␣HOOVER␣MASS:', T40, G25.16	1	Barostat mass.
VG_P_FIX	'␣HOOVER␣PRESSURE:', T40, G25.16	1	Barostat pressure.
VG_SCALERC	'␣SCALE␣RCUT␣WITH␣BOXLEN:', T40, L1	1	Scale cut-off radius.
VG_QE	'␣GE␣EXCHANGE␣MASS:', T40, G25.16	1	Gibbs ensemble exchange mass.
VG_FRAC_ETA	'␣GE␣FRACTIONAL␣ETA:', T40, G25.16	1	Fractional distance increase.

VG_NU_POS	'␣GE␣NU␣START␣POSITION:', T40, G25.16	1	Exchange start position.
VG_NU_VEL	'␣GE␣NU␣START␣VELOCITY:', T40, G25.16	1	Exchange start velocity.
VG_EXCSTART	'␣GE␣EXCHANGE␣START:', T40, I6	1	First exchange step.
VG_EXCBREAK	'␣GE␣EXCHANGE␣BREAK:', T40, I6	1	Break between exchanges.
USE_DHG	'␣USE␣DENSITY␣HISTOGRAM:', T40, L1	1	Create density histograms.

If USE_DHG is true, then the following values are written:

Variable	Type	Count	Description
DHG_START	'␣DHG␣-␣START␣DENSITY:', T40, G25.16	1	Start density.
DHG_END	'␣DHG␣-␣END␣DENSITY:', T40, G25.16	1	End density.
DHG_INTS	'␣DHG␣-␣INTERVALS:', T40, I6	1	Number of intervals.

All masses are in units of $\mathrm{kg\,m^2}$, temperatures in K, pressures in Pa and distances in m.

B.1.7. Additional Gibbs Ensemble Values

The creation of an exchange particle in the destination box needs the knowledge of several parameters of the input program. The equilibrium geometry, rotation axes and normal mode vectors are stored in the section $GEOM: of the .PAR file.

For each molecular species I:

Variable	Type	Count	Description
NOAT(I)	'␣NUMBER␣OF␣SITES:', T40, I10	1	Number of sites.

For each site:

Variable	Type	Count	Description
POS	'␣ATOM␣TAG:', T40, I10	1	Site tag.
REX(J)	'␣REX:', T40, G25.16	1	x-Position.
REY(J)	'␣REY:', T40, G25.16	1	y-Position.
REZ(J)	'␣REZ:', T40, G25.16	1	z-Position.

The distances are in units of m.

Variable	Type	Count	Description
COUNT	' ROTATIONS:', T40, I10	1	Number of rotation axes.

For each rotation axis[2]:

Variable	Type	Count	Description
AXIS(1)	' ROTATION AXIS X:', T40, G25.16	1	x-Coordinate.
AXIS(2)	' ROTATION AXIS Y:', T40, G25.16	1	y-Coordinate.
AXIS(3)	' ROTATION AXIS Z:', T40, G25.16	1	z-Coordinate.

Variable	Type	Count	Description
COUNT	' NORMAL MODES:', T40, I10	1	Number of normal modes.

For each normal mode:

Variable	Type	Count	Description
OMEGA	' OMEGA:', T40, G25.16	1	Frequency [s^{-1}].
ATOMS	' REAL ATOMS:', T40, I10	1	Number of real atoms.
MAXDISPLX	' MAX DISPLACEMENT X:', T40, G25.16	1	Maximum x-displacement.
MAXDISPLY	' MAX DISPLACEMENT Y:', T40, G25.16	1	Maximum y-displacement.
MAXDISPLZ	' MAX DISPLACEMENT Z:', T40, G25.16	1	Maximum z-displacement.

B.1.8. Correlation Manager

The correlation manager was extended to support the correlation of data structures of any dimension. The old version was limited to arrays. It is now possible to correlate tensors and to use the same correlation function more than once. The latter was necessary to use the local density correlation function at different densities.

An extended CORR_DATA data structure was introduced to store the input values. The updated simulation program is able to read the old and new format.

[2] Axis through the center of mass and the given point.

Type	:: **Variable**
TYPE(SM_CORRMGR_DATA), POINTER	:: NEXT
INTEGER	:: NLABEL
INTEGER	:: MAXDATA
INTEGER	:: NCSTEP
INTEGER	:: NTCOR
INTEGER	:: IOR
INTEGER	:: NTIMOR
INTEGER	:: FULLUP
INTEGER	:: DIM
INTEGER	:: AKTIOR
INTEGER	:: NINCOR
INTEGER	:: SETIOR
INTEGER	:: BEFIOR
INTEGER	:: ART_ID
CHARACTER(40)	:: FILENAME
Old Format	
REAL(SMR), POINTER, DIMENSION(:,:)	:: STORX,STORY,STORZ
REAL(SMR), POINTER, DIMENSION(:)	:: VACF
REAL(SMR), POINTER, DIMENSION(:)	:: ANORM
INTEGER, POINTER, DIMENSION(:)	:: DIIOR
INTEGER, POINTER, DIMENSION(:)	:: GLOB_T
New Format	
INTEGER	:: STORDIM
REAL(SMR), POINTER, DIMENSION(:,:,:)	:: STORAGE
REAL(SMR), POINTER, DIMENSION(:)	:: VACF
REAL(SMR), POINTER, DIMENSION(:)	:: ANORM
INTEGER, POINTER, DIMENSION(:)	:: DIIOR
INTEGER, POINTER, DIMENSION(:)	:: GLOB_T
INTEGER	:: DATADIM
REAL(SMR), POINTER, DIMENSION(:)	:: DATA
INTEGER	:: FUNCNR

Table B.5.: CORR_DATA Structure.

The additional DATA array can be used to store additional data assigned to each individual correlation function.

The new correlation function selection looks like this:

Subroutine: INPUTCORRDATA

```
CHOOSE CORRELATION FUNCTION:
****************************

 Self diffusion           ( 1)
```

```
 Shear viscosity flex      ( 2)
 Shear viscosity atom      ( 3)
 Shear viscosity rigid     ( 4)
 therm. conduct. flex      ( 5)
 therm. conduct. atom      ( 6)
 therm. conduct. rigid     ( 7)
 tot. relax. time (diag)   ( 8)
 tot. relax. time (offdiag)( 9)
 tau_1 for vec 1, CO or OD (10)
 tau_2 for vec 1, CO or OD (11)
 tau_1 for vec 2, OO or HD (12)
 tau_2 for vec 2, OO or HD (13)
 Local density correlation (14)

 NO CORRELATION FUNCTION   (0)

$CORFUNC:
```

The usual variables are read, as documented in reference [11].

The local density correlation functions read some additional variables:

```
 Radius of the sphere around a molecule (in Ang)?
$DATA1
```

Variable set: R, stored in DATA(1)

```
 Number of local g-functions?
```

Variable set: NRGFUNCS, stored in DATA(2)

NRGFUNCS times:

```
 g-func density (mol/m**3)?
```

The molar density is converted to the nearest number of molecules in the specified sphere.

Variable set: GFCMINDENS, stored in DATA(I+2)

B.2. Simulation Program

B.2.1. Molecular Constraints

The convergence conditions for the SHAKE algorithm were not strict enough. Long time simulations showed a small drift to higher energies. This was fixed with stronger conditions and a different algorithm. The one used in SimMol was slightly different to the one given in

microfiche 8 of Allen & Tildesley[3]. A RATTLE[5] version for the velocity Verlet algorithm was also developed. The code is still in the program files.

The derivation of the constraint equations will be shown in this chapter.

The SHAKE algorithm was developed by Ryckaert *et al.* [69] for chain molecules.

The complete algorithm for each CO_2 molecule is:

```
DONE = .FALSE.

DO I = 1,MAXITER
   IF (DONE) EXIT

   DONE = .TRUE.
   CALL SHAKEGEOM(...,DONE)
   CALL SHAKEBOND(1,2,...,DONE)
   CALL SHAKEBOND(2,3,...,DONE)
END DO

IF (.NOT.DONE) THEN
   STOP 'Too many constraints iterations (SHAKE)!'
END IF
```

The maximum number of iterations (`MAXITER`) is 30. The tolerance for the geometry is $|\sigma|^2 \leq d_{CO}10^{-6}$ (equation (B.23)) and for the bond length $|\sigma| \leq d_{CO}^2 \cdot 10^{-6}$ (equation (B.3)).

Rigid Bond

The SHAKE algorithm[3, p. 92] for one bond is shown here.

$$m_1 \ddot{\boldsymbol{q}}_1 = \boldsymbol{f}_1 + \boldsymbol{g}_1 \tag{B.1}$$
$$m_2 \ddot{\boldsymbol{q}}_2 = \boldsymbol{f}_2 + \boldsymbol{g}_2 \tag{B.2}$$

The Verlet algorithm displaced the positions $\boldsymbol{q}_i(t)$ to $\boldsymbol{q}_i(t + \delta t)$. $\boldsymbol{f}_i$ is the known force on the particle $\boldsymbol{q}_i$. $\boldsymbol{g}_i$ is an unknown constraint force which moves the particle to fulfill the constraint condition σ:

$$\sigma = |\boldsymbol{q}_{12}(t)|^2 - d_{12}^2 \equiv 0 \tag{B.3}$$
$$\boldsymbol{q}_{12} = \boldsymbol{q}_2 - \boldsymbol{q}_1 \tag{B.4}$$

Where d_{12} is the fixed bond length. The sum of the internal forces has to be zero:

$$\boldsymbol{g}_1 \equiv -\boldsymbol{g}_2 \tag{B.5}$$

The Lagrangian equations of motion are calculated:

$$\begin{aligned}
\mathcal{L} &= \boldsymbol{q}_{12}^2(t) - d_{12}^2 && \text{(B.6)}\\
\dot{\boldsymbol{p}} &\equiv \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} && \text{(B.7)}\\
\boldsymbol{g}_1 &= \frac{1}{2}\lambda \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}_1} && \text{(B.8)}\\
\boldsymbol{g}_2 &= \frac{1}{2}\lambda \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}_2} && \text{(B.9)}
\end{aligned}$$

Where λ is the Lagrangian multiplier. The prefactors $\frac{1}{2}$ simplify both equations to:

$$\begin{aligned}
\boldsymbol{g}_1 &= -\lambda\,(\boldsymbol{q}_2 - \boldsymbol{q}_1)\\
&= -\lambda\,\boldsymbol{q}_{12} && \text{(B.10)}\\
\boldsymbol{g}_2 &= \lambda\,(\boldsymbol{q}_2 - \boldsymbol{q}_1)\\
&= \lambda\,\boldsymbol{q}_{12} && \text{(B.11)}
\end{aligned}$$

Adding the constraint forces to the Verlet integrator and calculating the bond distance at the time $t + \delta t$ gives:

$$\begin{aligned}
\boldsymbol{q}_1(t+\delta t) &= \boldsymbol{q}_1'(t+\delta t) - \frac{\lambda\,\boldsymbol{q}_{12}\,\delta t^2}{m_1} && \text{(B.12)}\\
\boldsymbol{q}_2(t+\delta t) &= \boldsymbol{q}_2'(t+\delta t) + \frac{\lambda\,\boldsymbol{q}_{12}\,\delta t^2}{m_2} && \text{(B.13)}\\
\boldsymbol{q}_{12}(t+\delta t) &= \boldsymbol{q}_{12}'(t+\delta t) + \frac{\lambda\,\boldsymbol{q}_{12}\,\delta t^2\,(m_1+m_2)}{m_1 m_2} && \text{(B.14)}
\end{aligned}$$

$\boldsymbol{q}_{12}$ is inserted into the constraint condition (B.3) for the following time step, the quadratic terms are expanded and all quadratic values of λ are ignored. The final value of λ is:

$$\left(\boldsymbol{q}_{12}'(t+\delta t) + \frac{\lambda\,\boldsymbol{q}_{12}\delta t^2\,(m_1+m_2)}{m_1\,m_2}\right)^2 - d_{12}^2 = 0 \tag{B.15}$$

$$\lambda = -\frac{m_1\, m_2\, (\boldsymbol{q}'^2_{12}(t+\delta t) - d^2_{12})}{2\,(\boldsymbol{q}'_{12}(t+\delta t)\cdot \boldsymbol{q}_{12}(t))\,\delta t^2 (m_1+m_2)} \tag{B.16}$$

The forces are known from equations (B.10) and (B.11). The corrected positions are calculated by:

$$\boldsymbol{q}_1(t+\delta t) = \boldsymbol{q}'_1(t+\delta t) + \frac{\boldsymbol{g}_1}{m_1}\,\delta t^2 \tag{B.17}$$

$$\boldsymbol{q}_2(t+\delta t) = \boldsymbol{q}'_2(t+\delta t) + \frac{\boldsymbol{g}_2}{m_2}\,\delta t^2 \tag{B.18}$$

The corrected velocities are:

$$\begin{aligned} \dot{\boldsymbol{q}}_i(t) &= \frac{\boldsymbol{q}_i(t+\delta t) - \boldsymbol{q}_i(t-\delta t)}{2\,\delta t} \\ &= \frac{\boldsymbol{q}'_i(t+\delta t) + \Delta\boldsymbol{r}(t+\delta t) - \boldsymbol{q}_i(t-\delta t)}{2\,\delta t} \\ &= \dot{\boldsymbol{q}}'(t+\delta t) + \frac{\Delta\boldsymbol{r}(t+\delta t)}{2\,\delta t} \end{aligned} \tag{B.19}$$

The intramolecular virial of this pair interaction is:

$$\mathcal{W}_{12} = \frac{1}{3}\,(\boldsymbol{g}_1(t)\cdot\boldsymbol{q}_{12}(t)) \tag{B.20}$$

$$= -\frac{1}{3}\,\lambda\,(\boldsymbol{q}_{12}(t)\cdot\boldsymbol{q}_{12}(t)) \tag{B.21}$$

$$= -\frac{1}{3}\,\lambda\, d^2_{12} \tag{B.22}$$

Geometry

CO_2 is a linear three atomic molecule with the carbon at the center. The linear vectorial constraint is:

$$\boldsymbol{\sigma} = -\boldsymbol{q}_1 + 2\,\boldsymbol{q}_2 - \boldsymbol{q}_3 \equiv \boldsymbol{0} \tag{B.23}$$

The internal forces are:

$$\begin{aligned} \boldsymbol{g}_1 &= \frac{1}{2}\lambda\frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{q}_1} && \text{(B.24)}\\ &= -\frac{\lambda}{2} \\ \boldsymbol{g}_2 &= \frac{1}{2}\lambda\frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{q}_2} && \text{(B.25)}\\ &= \lambda \\ \boldsymbol{g}_3 &= \frac{1}{2}\lambda\frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{q}_3} && \text{(B.26)}\\ &= -\frac{\lambda}{2} \end{aligned}$$

The corrected positions are given by:

$$\begin{aligned} \boldsymbol{q}_1(t+\delta t) &= \boldsymbol{q}_1'(t+\delta t) - \frac{\lambda\,\delta t^2}{2\,m_O} && \text{(B.27)}\\ \boldsymbol{q}_2(t+\delta t) &= \boldsymbol{q}_2'(t+\delta t) + \frac{\lambda\,\delta t^2}{m_C} && \text{(B.28)}\\ \boldsymbol{q}_3(t+\delta t) &= \boldsymbol{q}_3'(t+\delta t) - \frac{\lambda\,\delta t^2}{2\,m_O} && \text{(B.29)} \end{aligned}$$

These positions are inserted into equation (B.23) and the value for λ is:

$$\lambda = \frac{\boldsymbol{q}_1'(t+\delta t) - 2\boldsymbol{q}_2'(t+\delta t) + \boldsymbol{q}_3'(t+\delta t)\,m_O\,m_C}{\delta t^2\,(m_C + 2m_O)} \qquad \text{(B.30)}$$

The position and velocity correction is done in the same way as in the bond example above. There is no contribution to the virial, the sum of the individual terms is by definition zero.

B.2.2. Restart Files

The extended parameters are stored in the usual `.PAR` files. Together with the `.XYZ` file, simulations can be restarted or continued. The accumulator values were initialized after the equilibration. All continued simulations inherit the previous values. To reset all accumulator values, a flag was added to the `.PAR` file. If the flag is zero then previous values are inherited. If the flag is not zero then all accumulators are reset at the beginning and new independent values are used to calculate the averages.

Variable	Type
RESTART_NAV_RESET	' NAVRESET: . NE. 0. -> RESET ACCS', G25.16

If the accumulator reset is once activated then it will be enabled for later runs.

B.3. Auswert Program

The Auswert program was extended by the menu points (6) to (8).

```
SIMMOL AUSWERTE PROGRAMM
================================
(1) G-FUNKTION
(2) KORRELATIONS FUNKTIONEN
(3) EINSTEIN DIFFUSION
(4) KORRELATIONS TABELLE
(5) XMOL FILM
(6) THERMODYNAMISCHE DATEN
(7) Local Density Distribution
(8) Partial Radial Distribution
--------------------------------
(9) OFFLINE CORRELATION
--------------------------------
(0) EXIT
```

B.3.1. g-Function

The program asks first if a Gnuplot file should be created.

```
Generate Gnuplot file? (y/n)
```

If yes was selected, the Gnuplot file "*stem*-GFCS.DAT" is created. If this file is opened with Gnuplot, a Postscript file containing all site-site g-functions is created.

B.3.2. Correlation Functions

The program asks first if a Gnuplot file should be created.

```
Generate Gnuplot file? (y/n)
```

If yes was selected, the Gnuplot file "*stem*.DAT" is created. If this file is opened with Gnuplot, Postscript files, containing each a plot of a correlation function, are created.

If a local density correlation function was used, the following question is asked.

```
Analyze local density? (y/n)
```

Local Density

The "*stem*-PROB.DAT" file contains two Gnuplot graphs showing the "local density neighbor probability" and the "local density probability".

The "*stem*-GFUNC.DAT" file contains the normal carbon-carbon g-function.

The "*stem*-LGFUNC*number*.DAT" file the local g-functions for a selected density.

The statistical density values are written to the screen:

```
<density>              = 14336.30     mol/m^3
<dens(l)/dens>         = 1.246635
<dens(l)-dens>         = 2836.299     mol/m^3
<(dens(l)-dens)/dens> = .2466347
```

B.3.3. Thermodynamic Data

Prints the data which was shown at the end of the simulation.

B.3.4. Local Density Distribution

```
Local density calculation
```

First the program asks for the r_{ck} value in Å.

```
Maximal radius of the sphere: 28 A
```

Then how many times the box should be concatenated in each dimension.

```
Number of boxes per dimension:
```

The three density regions are calculated. The results are stored in different XMol files. To simplify the graphs, only the center of mass position (carbon) is saved. In the low density region, the carbon is replaced by a sulfur and in the high density region by an oxygen.

File	Description
LDMAP.xyz	All regions
LDMAP-lo.xyz	Low density region
LDMAP-av.xyz	Average density region
LDMAP-hi.xyz	High density region
LDMAP-lohi.xyz	Low and high density regions

Table B.6.: Local Density Map Output Files.

The file LDMAP.DAT contains the x-y values of a density probability histogram. The x values are relative to the bulk density, showing the local density enhancement.

Finally the region values are printed:

```
N_LOW:     231
N_AVERAGE: 399
N_HIGH:    701
```

B.3.5. Partial Radial Distribution

First the program asks for the r_{ck} value in Å.

```
Sphere Radius for hi/av/lo (Angstrom)?
```

All available snapshot files are loaded to generate the g-functions. The three total g-functions and a graph with all partial g-functions is stored in the "*stem*-PGFC.DAT" Gnuplot file. The raw x-y values are stored in the files "*stem*.*[*lo|av|hi*]*.TGFC" and "*stem*.*[*lo|av|hi*]*-*[*lo|av|hi*]*.PGFC".

The g-functions are normalized to their own values, therefore at longer distances, they will be one, equal to ideal gas values. It is possible to normalize the values to the bulk density by changing the DENSMODE variable in the subroutine OUTPGFUNC of auswert.f90.

The final screen output shows the average number of molecules in each density zone and the number of snapshots which were used.

```
LDENS (particles):    5.30000000000000        6.10000000000000    8.90000000000000
Number of samplings:           10
Lo/Av/Hi:    243   377   709
```

C. Publications

1. E. H. Hardy, M. G. Müller, P. S. Vogt, C. Bratschi, B. Kirchner, H. Huber, D. J. Searles "How approximate is the experimental evaluation of quadrupole coupling constants in liquids? A novel computational study" *J. Chem. Phys.* **119** 6184-6193 (2003)

2. M. G. Müller, E. H. Hardy, P. S. Vogt, C. Bratschi, B. Kirchner, H. Huber, D. J. Searles "Calculation of the deuteron quadrupole relaxation rate in a mixture of water and dimethyl sulfoxide" *J. Am. Chem. Soc.* **126** 4704-4710 (2004)

3. M. Venkatraj, C. Bratschi, H. Huber, R. J. Gdanitz "Monte Carlo simulations of vapor-liquid equilibria of neon using an accurate ab initio pair potential" *Fluid Phase Equilibria* **218** 285-289 (2004)

A poster was presented at the "39th Symposium on Theoretical Chemistry 2003" (STC 2003) from 28 September to 2 October 2003 in Gwatt (Lake Thun, Switzerland).

Bibliography

[1] J. E. Adams and A. Siavosh-Haghighi "Rotational relaxation in supercritical CO_2" *J. Phys. Chem. B* **106** 7973–7980 (2002)

[2] B. J. Alder and T. E. Wainwright "Phase transition for a hard sphere system" *J. Chem. Phys.* **27** 1208–1209 (1957)

[3] M. P. Allen and D. J. Tildesley *Computer Simulations of Liquids* Clarendon Press, Oxford (1987)

[4] H. C. Andersen "Molecular dynamics simulations at constant pressure and/or temperature" *J. Chem. Phys.* **72** 2384–2393 (1980)

[5] H. C. Andersen "RATTLE - A velocity version of the SHAKE algorithm for molecular-dynamics calculations" *J. Comput. Phys.* **52** 24–34 (1983)

[6] S. Angus, B. Armstrong and K. M. de Reuck, editors *Carbon Dioxide: International Thermodynamic Tables of the Fluid State - 3* Pergamon Press, New York (1976)

[7] P. W. Atkins *Physikalische Chemie* VCH, Weinheim 5th edition (1996)

[8] P. F. Batcho and T. Schlick "Special stability advantages of position-Verlet over velocity-Verlet in multiple-time step integration" *J. Chem. Phys.* **115** 4019–4029 (2001)

[9] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola and J. R. Haak "Molecular dynamics with coupling to an external bath" *J. Chem. Phys.* **81** 3684–3690 (1984)

[10] S. Bock, E. Bich and E. Vogel "A new intermolecular potential energy surface for carbon dioxide from ab initio calculations" *Chem. Phys.* **257** 147–156 (2000)

[11] C. Bratschi *Das Molekül Dynamik Simulationsprogramm SimMol* Master's thesis University of Basel (2002) Diplomarbeit

[12] R. Bukowski, J. Sadlej, B. Jeziorski, P. Jankowski, K. Szalewicz, S. A. Kucharski, H. L. Williams and B. M. Rice "Intermolecular potential of carbon dioxide dimer from symmetry-adapted perturbation theory" *J. Chem. Phys.* **110** 3785–3803 (1999)

[13] R. Car and M. Parrinello "Unified Approach for Molecular Dynamics and Density-Functional Theory" *Phys. Rev. Lett.* **55** 2471–2474 (1985)

[14] K. Cho, J. D. Joannopoulos and L. Kleinman "Constant-temperature molecular dynamics with momentum conservation" *Phys. Rev. E* **47** 3145–3151 (1993)

[15] G. Ciccotti, G. J. Martyna, S. Melchionna and M. E. Tuckerman "Constrained isothermal-isobaric molecular dynamics with full atomic virial" *J. Phys. Chem. B* **105** 6710–6715 (2001)

[16] P. Cipriani, M. Nardone and F. P. Ricci "Neutron diffraction measurements on CO_2 in both undercritical and supecritical states" *Physica B* **241** 940–946 (1997)

[17] R. F. Cracknell, D. Nicholson, N. G. Parsonage and H. Evans "Rotational insertion bias: a novel method for simulating dense phases of structured particles, with particular application to water." *Mol. Phys.* **71** 931–943 (1990)

[18] Z. H. Duan, N. Møller and J. H. Weare "Gibbs Ensemble Simulations of Vapor/Liquid Equilibrium Using the Flexible RWK2 Water Potential" *J. Phys. Chem. B* **108** 20303–20309 (2004)

[19] T. H. Dunning "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen" *J. Chem. Phys.* **90** 1007–1023 (1989)

[20] R. Eckhardt "Stan Ulam, John von Neumann, and the Monte Carlo method" *Los Alamos Science* **Special Issue (15)** 131–137 (1987)

[21] D. J. Evans and B. L. Holian "The Nose-Hoover thermostat" *J. Chem. Phys.* **83** 4069–4074 (1985)

[22] P. W. Fowler, Z. Kisiel and A. Lotnikow "VIBCA - Vibrational Calculations Program" `http://info.ifpan.edu.pl/~kisiel/prospe.htm` (1999)

[23] D. Frenkel and A. J. C. Ladd "New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres." *J. Chem. Phys.* **81** 3188–3193 (1984)

[24] D. Frenkel and B. Smit *Understanding Molecular Simulation - From Algorithms to Applications* Academic Press, San Diego (2002)

[25] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, N. Rega, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts,

R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle and J. A. Pople *Gaussian 98, Revision A.11.3* Gaussian, Inc. Pittsburgh PA (2002)

[26] G. Goodyear, M. W. Maddox and S. C. Tucker "Correlation between local and long-range structure in compressible supercritical Lennard-Jones fluids: State-point dependence" *J. Phys. Chem. B* **104** 6258–6265 (2000)

[27] G. Goodyear, M. W. Maddox and S. C. Tucker "The correlation between local and long-range structure in compressible supercritical fluids" *J. Chem. Phys.* **112** 10327–10339 (2000)

[28] G. Goodyear and S. C. Tucker "What causes the vibrational lifetime plateau in supercritical fluids?" *J. Chem. Phys.* **110** 3643–3646 (1999)

[29] E. Hairer, C. Lubich and G. Wanner "Geometric numerical integration illustrated by the Störmer-Verlet method" *Acta Numerica* **12** 399–450 (2003)

[30] D. J. Hardy and D. I. Okunbor "Qualitative study of the symplectic Störmer-Verlet integrator" *J. Chem. Phys.* **102** 8978–8982 (1995)

[31] J. G. Harris and K. H. Yung "Carbon Dioxide's Liquid-Vapor Coexistence Curve and Critical Properties As Predicted by a Simple Molecular Model" *J. Phys. Chem.* **99** 12021–12024 (1995)

[32] S. C. Harvey, R. K. Z. Tan and T. E. Cheatham "The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition" *J. Comput. Chem.* **19** 726–740 (1998)

[33] R. Hentschke, T. Bast, E. Aydt and M. Kotelyanskii "Gibbs-Ensemble Molecular Dynamics: A New Method for Simulations Involving Particle Exchange" *J. Mol. Model.* **2** 319–326 (1996)

[34] B. L. Holian, A. J. De Groot, W. G. Hoover and C. G. Hoover "Time-reversible equilibrium and nonequilibrium isothermal-isobaric simulations with centered-difference Stoermer algorithms" *Phys. Rev. A* **41** 4552–4553 (1990)

[35] B. L. Holian, A. F. Voter and R. Ravelo "Thermostatted molecular dynamics: How to avoid the Toda demon hidden in Nosé-Hoover dynamics" *Phys. Rev. E* **52** 2338–2347 (1995)

[36] W. G. Hoover "Canonical dynamics: Equilibrium phase-space distributions" *Phys. Rev. A* **31** 1695–1697 (1985)

[37] W. G. Hoover "Constant-pressure equations of motion" *Phys. Rev. A* **34** 2499–2500 (1986)

[38] W. G. Hoover "Liouville's theorems, Gibbs' entropy, and multifractal distributions for nonequilibrium steady states" *J. Chem. Phys.* **109** 4164–4170 (1998)

[39] W. G. Hoover and C. G. Hoover "Links between microscopic and macroscopic fluid mechanics" *Mol. Phys.* **101** 1559–1573 (2003)

[40] J. M. Ilnytskyi and M. R. Wilson "A domain decomposition molecular dynamics program for the simulation of flexible molecules of spherically-symmetrical and nonspherical sites. II. Extension to NVT and NPT ensembles" *Comput. Phys. Comm.* **148** 43–58 (2002)

[41] G. Kalibaeva, M. Ferrario and G. Ciccotti "Constant pressure-constant temperature molecular dynamics: a correct constrained NPT ensemble using the molecular virial" *Mol. Phys.* **101** 765–778 (2003)

[42] G. R. Kneller and T. Mülders "Nosé-Andersen dynamics of partially rigid molecules: Coupling all degrees of freedom to heat and pressure baths" *Phys. Rev. E* **54** 6825–6837 (1996)

[43] M. J. Kotelyanskii and R. Hentschke "Gibbs-ensemble molecular dynamics: Liquid-gas equilibrium in a Lennard-Jones system" *Phys. Rev. E* **51** 5116–5119 (1995)

[44] V. Kräutler, W. F. van Gunsteren and P. H. Hünenberger "A fast SHAKE: Algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations" *J. Comput. Chem.* **22** 501–508 (2001)

[45] LAM Team "LAM/MPI" `http://www.lam-mpi.org/`

[46] M. W. Maddox, G. Goodyear and S. C. Tucker "Effect of critical slowing down on local-density dynamics" *J. Phys. Chem. B* **104** 6266–6270 (2000)

[47] M. W. Maddox, G. Goodyear and S. C. Tucker "Origins of atom-centered local density enhancements in compressible supercritical fluids" *J. Phys. Chem. B* **104** 6248–6257 (2000)

[48] H. L. Martinez, R. Ravi and S. C. Tucker "Characterization of solvent clusters in a supercritical Lennard-Jones fluid" *J. Chem. Phys.* **104** 1067–1080 (1996)

[49] G. J. Martyna "Remarks on "Constant-temperature molecular dynamics with momentum conservation"" *Phys. Rev. E* **50** 3234–3236 (1994)

[50] G. J. Martyna, D. J. Tobias and M. L. Klein "Constant pressure molecular dynamics algorithms" *J. Chem. Phys.* **101** 4177–4189 (1994)

[51] MathWorld "Delta Function" `http://mathworld.wolfram.com/DeltaFunction.html`

[52] MathWorld "L'Hospital's Rule" `http://mathworld.wolfram.com/LHospitalsRule.html`

[53] MathWorld "Liouville's Phase Space Theorem" `http://mathworld.wolfram.com/LiouvillesPhaseSpaceTheorem.html`

[54] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller "Equation of State Calculations by Fast Computing Machines" *J. Chem. Phys.* **21** 1087–1092 (1953)

[55] M. G. Müller *Model Calculation of Deuteron Quadrupole Coupling Constants and Quadrupole Relaxation Rates in Aqueous Systems* Ph.D. thesis Universität Basel (2003)

[56] K. Nishikawa, H. Ochiai, K. Saitow and T. Morita "Static inhomogeneity of supercritical ethylene studied by small-angle X-ray scattering" *Chem. Phys.* **286** 421–430 (2003)

[57] S. Nosé "A molecular dynamics method for simulations in the canonical ensemble" *Mol. Phys.* **52** 255–268 (1984)

[58] S. Nosé "A unified formulation of the constant temperature molecular dynamics methods" *J. Chem. Phys.* **81** 511–519 (1984)

[59] S. Okazaki, M. Matsumoto, I. Okada, K. Maeda and Y. Kataoka "Density dependence of rotational relaxation of supercritical CF_3H" *J. Chem. Phys.* **103** 8594–8601 (1995)

[60] B. J. Palmer and C. Lo "Molecular dynamics implementation of the Gibbs ensemble calculation" *J. Chem. Phys.* **101** 10899–10907 (1994)

[61] A. Z. Panagiotopoulos "Direct determination of phase coexistence properties of fluids by Monte Carlo simulations in a new ensemble" *Mol. Phys.* **61** 813–826 (1987)

[62] A. Z. Panagiotopoulos "Molecular simulation of phase coexistence: Finite-size effects and determination of critical parameters for two- and three- dimensional Lennard-Jones fluids" *Int. J. Thermophys.* **15** 1057 (1994)

[63] A. Z. Panagiotopoulos "Monte Carlo methods for phase equilibria of fluids" *J. Phys.: Condens. Matter* **12** R25–R52 (2000)

[64] M. Parrinello and A. Rahman "Crystal Structure and Pair Potentials: A Molecular-Dynamics Study" *Phys. Rev. Lett.* **45** 1196–1199 (1980)

[65] M. Poliakoff and P. King "Phenomenal fluids" *Nature* **412** 125–125 (2001)

[66] A. Rahman "Correlations in motion of atoms in liquid argon" *Phys. Rev. A* **136** A405–A411 (1964)

[67] J. D. Ramshaw "Remarks on non-Hamiltonian statistical mechanics" *Europhys. Lett.* **59** 319–323 (2002)

[68] J. P. Ryckaert and G. Ciccotti "Introduction of Andersen's demon in the molecular dynamics of systems with constraints" *J. Chem. Phys.* **78** 7368–7374 (1983)

[69] J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen "Numerical-Integration of Cartesian

Equations of Motion of a System with Constraints - Molecular-Dynamics of n-Alkanes" *J. Comp. Phys.* **23** 327–341 (1977)

[70] K. Saitow, K. Otake, H. Nakayama, K. Ishii and K. Nishikawa "Local density enhancement in neat supercritical fluid due to attractive intermolecular interactions" *Chem. Phys. Lett.* **368** 209–214 (2003)

[71] J. M. H. Sengers, J. Straub and M. Vicentin "Coexistence curves of CO_2, N_2O, and $CClF_3$ in the Critical Region" *J. Chem. Phys.* **54** 5034–5050 (1971)

[72] V. M. Shmonov and K. I. Shmulovich "Molal volumes and equation of state of CO_2 at temperatures from 100 to 1000°C and pressures from 2000 to 10000 bars" *Doklady Akad. Nauk SSSR* **217** 935–938 (1974)

[73] R. D. Skeel, G. H. Zhang and T. Schlick "A family of symplectic integrators: Stability, accuracy, and molecular dynamics applications" *SIAM J. Sci. Comput.* **18** 203–222 (1997)

[74] SourceForge.net "Jmol" `http://jmol.sourceforge.net/`

[75] SourceForge.net "OpenBabel" `http://openbabel.sourceforge.net/`

[76] G. Steinebrunner *Molekül̈dynamiksimulation zur Berechnung von Eigenschaften des fluiden CO_2* Ph.D. thesis Universität Basel (1998)

[77] G. Steinebrunner, A. J. Dyson, B. Kirchner and H. Huber "Structural and thermodynamic properties of fluid carbon dioxide from a new ab initio potential energy surface" *J. Chem. Phys.* **109** 3153–3160 (1998)

[78] H. A. Stern "Simple algorithm for isothermal-isobaric molecular dynamics" *J. Comput. Chem.* **25** 749–761 (2004)

[79] C. Störmer "Sur les trajectoires des corpuscles électrisé" *Arch. Sci. Phys. Nat.* **24** 5–18, 113–158, 221–247 (1907)

[80] W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson "A computer-simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules: Application to small water clusters" *J. Chem. Phys.* **76** 637–649 (1982)

[81] K. T. Tang and J. P. Toennies "An improved simple-model for the van der Waals potential based on universal damping functions for the dispersion coefficients" *J. Chem. Phys.* **80** 3726–3741 (1984)

[82] I. G. Tironi, R. M. Brunne and W. F. van Gunsteren "On the relative merits of flexible versus rigid models for use in computer simulations of molecular liquids" *Chem. Phys. Lett.* **250** 19–24 (1996)

[83] S. Toxvaerd "Molecular dynamics at constant temperature and pressure" *Phys. Rev. E* **47** 343–350 (1993)

[84] S. C. Tucker "Solvent density inhomogeneities in supercritical fluids" *Chem. Rev.* **99** 391–418 (1999)

[85] S. C. Tucker and M. W. Maddox "The effect of solvent density inhomogeneities on solute dynamics in supercritical fluids: A theoretical perspective" *J. Phys. Chem. B* **102** 2437–2453 (1998)

[86] M. E. Tuckerman, Y. Liu, G. Ciccotti and G. J. Martyna "Non-Hamiltonian molecular dynamics: Generalizing Hamiltonian phase space principles to non-Hamiltonian systems" *J. Chem. Phys.* **115** 1678–1702 (2001)

[87] M. E. Tuckerman and G. J. Martyna "Understanding Modern Molecular Dynamics: Techniques and Applications" *J. Phys. Chem. B* **104** 159–178 (2000)

[88] M. E. Tuckerman, C. J. Mundy and G. J. Martyna "On the classical statistical mechanics of non-Hamiltonian systems" *Europhys. Lett.* **42** 149–155 (1999)

[89] R. S. Urdahl, D. J. Myers, K. D. Rector, P. H. Davis, B. J. Cherayil and M. D. Fayer "Vibrational lifetimes and vibrational line positions in polyatomic supercritical fluids near the critical point" *J. Chem. Phys.* **107** 3747–3757 (1997)

[90] R. S. Urdahl, K. D. Rector, D. J. Myers, P. H. Davis and M. D. Fayer "Vibrational relaxation of a polyatomic solute in a polyatomic supercritical fluid near the critical point" *J. Chem. Phys.* **105** 8973–8976 (1996)

[91] L. Verlet "Computer "Experiments"on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules" *Phys. Rev.* **159** 98–103 (1967)

[92] P. S. Vogt "Intramolecular CO_2 Potential Calculation"

[93] J. Vorholz, V. I. Harismiadis, B. Rumpf, A. Z. Panagiotopoulos and G. Maurer "Vapor + liquid equilibrium of water, carbon dioxide, and the binary system, water + carbon dioxide, from molecular simulation" *Fluid Phase Equilibria* **170** 203–234 (2000)

[94] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles and D. J. Evans "Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales" *Phys. Rev. Lett.* **89** 050601–1–050601–4 (2002)

[95] B. Widom "Some Topics in the Theory of Fluids" *J. Chem. Phys.* **39** 2808–2812 (1963)

[96] Wikipedia "Wikipedia article about I. Newton." `http://en.wikipedia.org/wiki/Newton%27s_laws_of_motion`

[97] L. V. Woodcock "Isothermal molecular dynamics calculations for liquid salts" *Chem. Phys. Lett.* **10** 257–261 (1971)

Acknowledgements

First, I would like to thank Prof. Dr. Hanspeter Huber for giving me the opportunity to perform a PhD thesis in his research group and for his guidance and advice.

I would like to thank Prof. Dr. Markus Meuwly for the acceptance of the co-referee position at the exam.

I would like to thank all group members for the discussion and help: Dr. Patrick S. Vogt, Dr. Edme H. Hardy, Marco Engeler, Dr. Muthusamy Venkatraj, Dr. Markus G. Müller, Inna Boychenko and Stefka G. Tsintsarska. And the members of Prof. Dr. Markus Meuwly's group for their interesting seminars.

Personally, I would like to thank Dr. Patrick S. Vogt for the introduction to Unix and Linux systems; Dr. David Nutt for the English language corrections; Dr. Debra J. Bernhardt for her help with the initial SimMol program code and the invitation to Brisbane during the diploma.

I would like to thank Esther Stalder and Daniela Tischhauser for their administrative work in the institute; Georg Holderied and Urs Buser for their support of the infrastructure and interesting discussions.

I would like to thank my parents for their patience and support during the last years.

Curriculum Vitae

Lebenslauf

Persönliche Daten

Christoph Bratschi
Felsenweg 16
4123 Allschwil

Tel.: 061 481 09 87
E-Mail: christoph@bratschi.net

Geb. am 25. 04. 1977 in Basel
Heimatort: Safnern (BE)
Ledig

Schulbildung

1984–1989	Primarschule in Allschwil
1989–1993	Progymnasium in Allschwil
1993–1997	Mathematisch-Naturwissenschaftliches Gymnasium (MNG) Basel Abschluss mit Matura Typus C

Studium

10/1997–2/2002	Chemie Studium an der Universität Basel Wahlfach Informatik Zwei Semester Tutor im Institut für Informatik Abschluss mit Diplom
3/2002–4/2005	Dissertation am Departement Chemie (Computational Chemistry) Computer Assistent

Basel, 21. März 2005

Während meiner Ausbildung habe ich Lehrveranstaltungen folgender Dozenten der Universität Basel besucht:

Prof. Dr. Helmar Burkhart, Prof. Dr. Edwin C. Constable, Prof. Dr. Catherine Constable-Housecroft, Prof. Dr. Bernhard Giese, Prof. Dr. Peter C. Hauser, Prof. Dr. Hanspeter Huber, Prof. Dr. Hans-Christoph Im Hof, Prof. Dr. Martin Jungen, Prof. Dr. Thomas A. Kaden, Prof. Dr. John Paul Maier, Prof. Dr. Wolfgang P. Meier, Prof. Dr. Margareta Neuburger-Zehnder, Prof. Dr. Michael Oehme, PD Dr. Olaf Schenk, Prof. Dr. Christian Schönenberger, Prof. Dr. Urs Séquin, Prof. Dr. Helmut Sigel, Prof. Dr. Peter Strazewski, Prof. Dr. Ludwig Tauscher, Prof. Dr. Dirk Trautmann, Prof. Dr. Christian Ullrich, Prof. Dr. Angelo Vedani, Prof. Dr. Hans-Jakob Wirz, Prof. Dr. Wolf-Dietrich Woggon